Journal of Multimedia

ISSN 1796-2048
Volume 9, Number 9, September 2014

Contents

REGULAR PAPERS

An Identity Authentication Protocol Based on SM2 and Fingerprint USBkey
Zhang Lihua, Dai Xichang, and Shen Youjin

Multiscale Recognition Algorithm for Eye Ground Texture Based on Fusion Threshold Equalization
Zhongsheng Qiu, Fangfang Zhang, and Yuan Zhou

An Improved Two-Dimensional OTSU Segmentation Method for Nvshu Character Image
Que Dong, Jiangqing Wang, and Yangguang Sun

Improved HT Object Detection Algorithm Based on Canny Edge Operator
Tingnian He, Xiaohong Li, and Yun Jiang

Image Signal Enhancement based on Fractional Differential Technologies
Zuo Lijuan, Bai Cuixia, Yang Yunfan, and Zhuo Fengqin

Remote Sensing Classification based on Improved Ant Colony Rules Mining Algorithm
Liu Shuying

Automatic Target Detection Method for High-Resolution Synthetic Aperture Radar Images
Xu Hui and Gu Hong

Emotion Model of Crowd in Mass Incidents
Yin Yanjun and Tang Weiqing

Color Remote-sensing Image Segmentation Based on Improved Region Filter
Hou Lei

1069
1075
1081
1089
1097
1105
1113
1120
1128
An Identity Authentication Protocol Based on SM2 and Fingerprint USBkey

Zhang Lihua
School of Software, East China Jiaotong University, Nanchang, China
School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai, China

Dai Xichang and Shen Youjin
School of Electrical and electronic engineering, East China Jiaotong University, Nanchang, China

Abstract—In order to solve network security problem, we need identity authentication protocol to ensure legal user’s authority. Multi-factor identity authentication protocols have more merits than the common identity authentication protocol. This paper analyzes the shortcomings of existing network identity authentication methods, and proposes a new authentication protocol based on SM2 and fingerprint USBkey. The proposed scheme which combines the fingerprint USBkey of fingerprint certificate with the National cipher algorithm SM2, constructs the multi-factor authentication model. And the scheme obtains many merits: 1) it adopts the challenge response authentication mechanism, and realizes the multi-factor mutual authentication. 2) it implements USBkey to verify the user and the remote server authentication by fingerprint features. 3) it can prevent eavesdropping attack, impersonation attack, replay attack and dos attack effectively. 4) it has better calculation and security performance than the existed schemes.

Index Terms—USBkey; SM2; Identity Authentication; Fingerprint

I. INTRODUCTION

The network identity authentication protocol is an important mean to solve the problem of increasingly serious network security problems. However, the common identity authentication schemes have many shortcomings. The static password authentication schemes are vulnerable to password guessing attack and dictionary attack, and passwords plaintext transfer are eavesdropped easily [2]. The dynamic password authentication’s time mechanism stepped out and synchronization events became complicated, because of the existence of time drifting [3-5]. The smart card authentication suffers smart card loss/replication attack [6-8]. A kind of identity authentication based on fingerprint characteristics is proposed by Ge Xiaomin et al [9], but in this scheme, the fingerprint characteristic is transmitted on the unsafe network, which will lead to the illegal invaders steal fingerprint image in the process of network transmission and conduct replay attack [10-12].

Because the existing network identity authentication has respective flaws and traits, scholars put forward some improved measures, that is mixing together static/dynamic password authentication, smart card certification, certification of USBkey, and biometric authentication methods, forming a multifactor authentication [13, 14]. But due to the complexity of the identity authentication and the diversity of network attacks, the authentication scheme for simple physical integration, cannot solve the problem of network identity authentication completely. The smart card authentication scheme based on ECC (Elliptic Curve Cryptography) and iris recognition was proposed by Shu-li Song et al [15], it is still difficult to resist the smart card loss/replication attack. As for the iris feature extraction is complex so that it is inconvenient for daily usage. The dynamic identity authentication scheme based on USBkey was proposed by Liu Huaila et al [16], because the generation of random numbers in the scheme has certain regularity, and USBkey still adopts a protective way of PIN number to control permission, leading to the security of the scheme is not high. In addition, for the security problems based on the commercial password, the State Cryptography Administration stipulates in 1 July 2011, the State Cryptography Administration stipulates that the commercial asymmetric cryptographic algorithm should use the SM2 elliptic curve cryptographic algorithms [17] what is released by the State Cryptography Administration. For the problems in network identity authentication and the existing research results. A fingerprint USBkey identity authentication scheme based on the national cipher algorithm SM2 is proposed in this paper. By putting a fingerprint authentication into the USBkey, using fingerprint instead of the PIN number to verify USBkey access, and the fingerprint is incorporated into the remote identity authentication for user. Through the SM2 asymmetric cryptographic algorithm to sign the interactive information in identity authentication, the security of data transmission is ensured. Adopting national cipher algorithm can make the scheme well applied in the business field.

This paper is organized as follows: In sect. 2, the SM2 and fingerprint USBkey is introduced. A novel identity authentication protocol is proposed in sect. 3, and the security analysis of the scheme is also demonstrated and discussed in sect. 4. Finally, some conclusions are drawn in sect. 5.
II. NATIONAL CIPHER ALGORITHM SM2 AND FINGERPRINT USBKEY

A. National Cipher Algorithm SM2

The identity authentication scheme proposed in this paper uses the national cipher algorithm SM2 [9] based on elliptic curve discrete logarithm problem. As public key cryptography algorithm, SM2 can complete the signature, key exchange and data encryption. The SM2 national cipher standard stipulated the signature, the calculation procedure of attestation and encryption, decryption calculation process. The national cipher algorithm SM2 has more clear advantages of security than the widely used cipher algorithm RSA based on factorization problem of large number, and it is much smaller in the same security, consequently, the calculation are faster, and the efficiency is higher. What's more, the SM2 applied in the commercial field could avoid backdoor attack hidden danger of foreign cipher algorithm procedure.

1) The National Cipher SM2 Signature Algorithm

Assuming the both sides of communication for A and B, A sends the message M to B, its national cipher SM2 signature algorithm is as follows:

(1) The elliptic curve is defined as $E(F_p)$, A is the base point that order is n on elliptic curve. The user A chooses the private key $d_A$ to calculate the public key $P_A = d_A G$. $H(\cdot)$ is the Hash function, $Z_A$ is the hash value about identifiable logo of User A, part of the elliptic curve system parameters and public key of User A.

(2) Signature: A calculates $\tilde{M} = Z_A | M$, then calculate the point on the elliptic curve i.e. $(x_1, y_1) = kG$ ($k$ is random number), finally, calculates $r = (e + x_1) \mod n$, $s = ((1 + d_A) \cdot (k - r \cdot d_A)) \mod n$. The signature pair for the message M is $(r, s)$. A sends $(r, s)$ and message M to B.

(3) Verification: After receiving $(r, s)$ and message M, B calculates $t = (r + s) \mod n$, $(x_1, y_1) = sG + tP_A$. Then B calculates $e' = H(Z_A | M)$, $R = (e', x_1) \mod n$. And B verifies whether $R$ is equal to $r$. If equal, the signature verification was passed, Otherwise, the validation fails.

2) The National Cipher SM2 Encryption Algorithm

Assuming the public and private key pair for B is $(d_B, P_B)$, A uses the public key of B to encrypt message M whose length is l, then A decrypts the encrypted cipher text. The process of encryption and decryption as follows:

(1) Encryption: A calculates $(x_2, y_2) = kP_B$, and calculates the cipher text $C = C_1 \mid C_2 \mid C_3$, in which

$$C_1 = kG(x_2, y_2), \quad C_2 = M \oplus t \quad (t = KDF(x_2 | y_2, l),$$

$KDF(\cdot)$ is Key derivation function),

$$C_3 = H(x_2 | M | y_2).$$

A sends C to B.

(2) Decryption: B calculates $d_B C_1 = (x_3, y_3)$, $t = KDF(x_3 | y_3, l)$, finally outputs plaintext $M = C_2 \oplus t$.

B. The Fingerprint USBkey

The scheme of fingerprint USBkey identity authentication, which combines SM2 cipher algorithm and fingerprint identification technology with USBkey to construct a new type of fingerprint USBkey. The fingerprint USBkey is a kind of USB interface equipment, which is a collection of fingerprint sensor, information security fingerprint processing chip, information security chip, and embedded COS, etc. The fingerprint sensor and the fingerprint processing chip construct the fingerprint processing module of fingerprint USBkey, and the information security chip with the embedded COS constructs the information security module. Its structure is shown in figure 1.

![Figure 1. The structure of fingerprint USBkey](image)

The fingerprint USBkey built-in SM2 cipher algorithm, stores user private key and digital certificate including template of fingerprint characteristic, etc. The fingerprint USBkey can be used for digital signature and encryption/decryption arithmetic. Because of the fingerprint USBkey stores the user key in the key storage area of USBkey, which is unable to export USBkey, all encryption/decryption and signature are carried out in USBkey, ensuring the safety of the private key.

We Use the fingerprint characteristic to improve the traditional USBkey with the method of PIN protection, to prevent the PIN code being stolen or being stolen by Trojan, and make USBkey form one-to-one correspondence relationship with the user entity identity, to realize USBkey entity authentication for user. Furthermore, we use the certificate from the user's fingerprint characteristics, to implement the remote service entity authentication for the user.

In order to make the user's fingerprint characteristic not to be tampered with or copied, we use the trusted certificate authority (CA) to generate digital certificate including the fingerprint characteristics with the user's public key bindings. And we promulgate the only corresponding certificate serial number i.e. $f_{sid}$. Upon the USBkey initialization, the digital certificate is solidified in USBkey, the fingerprint characteristic certificate contains the user's fingerprint characteristic $Temp$, which is used for the USBkey access authentication and for remote authentication. The specific process is as follows:

(1) Using fingerprint feature instead of PIN, to implement USBkey about user entity authentication. When user uses USBkey every time, the user must input the fingerprint firstly, then the user's fingerprint characteristic is generated by fingerprint processing chip,
and it is matched with the fingerprint characteristic Temp. Only the feature matching is successful, the user can access USBkey.

(2) Using fingerprint characteristic to implement the remote service entity authentication for the user. When the remote authenticating, the fingerprint feature matching is successful, releasing the sequence number \( f_u \) corresponding to the fingerprint characteristics for remote authentication.

III. PROTOCOL OF IDENTITY AUTHENTICATION

The fingerprint USBkey identity authentication protocol based on SM2 includes phase of user registration, remote authentication, password, fingerprint characteristics change and log off. The main body of the remote authentication includes user, fingerprint USBkey, the remote server and the access control center (ACC) of the remote server. The server provides a remote application server for User, ACC conducts authentication control for user access to the server. Only the legitimate authenticated users are allowed to access to the server. And ACC is responsible for authenticating and issuing the digital certificate that containing fingerprint characteristics.

A. Application for User Registration

Before the user registers to ACC, ACC uses SM2 algorithm to generate the key pair i.e \( (x_u, K_u) \) \((x_u \) is private key of ACC and \( K_u \) is its public key), and by using the SM3 algorithm to generate the Hash function \( h() \). ACC publicizes \( K_u \) and \( h() \) to all registered users. The user registration process is shown in figure 2.

(1) User requests to access server for ACC, and sends \((ID_i, PW_i)\) to ACC.

(2) After receiving \((ID_i, PW_i)\), ACC employs \( K_u \) in the message \( M_i \) to achieve the result of decryption i.e \( H_1, H_2 \) and \( Temp \), as well calculates \( H_2 = h(Temp) \). Then ACC verifies whether \( H_2 \) is equal to \( H_1 \). If it is equal, the digital certificate \( Cert_i \), including fingerprint is issued for user’s \( ID_i \). In addition, \( Cert_i = E_{K_u}(Temp|f_u)|K_u|ID_i|Sing_{K_u} \). Certificate including the encrypted fingerprint features and fingerprint sequence number \( f_u \), the user’s public key, the signature of the ACC etc. The ACC add user list \( U_i \), which includes \( Cert_i \), \( H_1 \), ACC sends \( Cert_i \) to USBkey.

(3) USBkey downloads \( Cert_i \) to digital certificate storage area of USBkey. By the form of encrypting fingerprint characteristic certificate, which can prevent the illegal user from being tampered with the fingerprint characteristics of USBkey effectively.

B. Remote Authentication Process

The remote authentication process is shown in figure 3.

The specific process as follows:

(1) User inputs \((ID_i, PW_i)\) into USBkey. After receiving \((ID_i, PW_i)\), USBkey calculates \( H_1 = h(ID_i, PW_i) \), and verifies whether \( H_1 \) is equal to \( H_1 \), if it is equal, ACC will calculate \( M_2 = E_{x_u}[E_{K_u}[n, h(ID_i, PW_i)]] \) (ure is random number ). Then ACC will send \( M_2 \) to USBkey of user.

(2) Fingerprint USBkey employs its own private key \( x_u \) and public key \( K_u \) of ACC to \( M_2 \) after receiving \( M_2 \) for obtaining \( h(ID_i, PW_i) \) and \( n \). Finger USBkey then verifies whether \( h(ID_i, PW_i) \) is equal to \( H_1 \) in fingerprint USBkey. If equal, the fingerprint USBkey pass the certification for ACC.

(3) Fingerprint USBkey prompts the user to input fingerprints, after the user inputs fingerprints, USBkey extracts fingerprint image feature and exploits fingerprint template in the certificate of fingerprint characteristic to compare with user’s fingerprint. If success, it indicates
that the user has access to the fingerprint USBkey, implementing the USBkey entity authentication for the user.

5) After passing entity authentication, USBkey calculates \( M_3 = E_k [E_y (n, c, f_d)] | ID \) (c is random number) and sends it to ACC.

6) After receiving \( M_1 \), ACC finds out sequence table \( U_i \) of users and obtains the public key \( K_s \) of USBkey via \( ID \), and employs its own private key \( x_i \) and \( K_s \) to decrypt \( M_3 \) for obtaining \( n \), \( c \) and \( f_d \). ACC then verifies whether \( n \) is equal to \( n \) and \( f_d \) is equal to \( f_d \). If equal, USBkey passed certification.

C. The Phase That Password and Fingerprint Characteristics Change and Log Off

If users need to change the new ID, password and fingerprint template, identity authentication has to be conducted among user, fingerprint USBkey and ACC/Server, the ID password and fingerprint template have to change after the identity authentication. The specific process is as follows:

1) The change of user ID and password: Inputting new (\( ID', PW' \)) into USBkey, and USBkey calculates \( H_i = h(ID', PW') \), \( M_4 = E_y (H_1, ID') | K_s | ID' \), afterwards, sending ACC the \( M_4 \) to update user's sequence table, and calculates the new digital certificate to fingerprint USBkey, updating the certificate.

2) The change of user fingerprint template: USBkey achieves user's fingerprint image to generate a new template \( Temp' \) of fingerprint characteristic, and calculates \( M_5 = E_y (H_1, H_2, Temp') | K_s | ID' \) to ACC. Depending on \( M_5 \), ACC calculates the new digital certificate containing the fingerprint characteristics and sends it to USBkey. USBkey downloads the digital certificate to the storage area of USBkey certificate.

IV. PERFORMANCE ANALYSIS

A. Security Analysis

For the identity authentication technology is proposed in this scheme, analyzing its ability about resist attacks and security.

1) The Security Identity Authentication Among Subjects

Analysis: The mutual authentication relationship among user, fingerprint USBkey and ACC/Server is shown in Figure 4.

![Figure 4. The mutual authentication relationship among subjects](image)

The figure 4 shows that the fingerprint USBkey and ACC/Server implement two-factor authentication for User, and fingerprint USBkey conducts the entity authentication for user, improving the security of the identity authentication. Utilizing the user's fingerprint characteristics and the random number achieves the two-way authentication between fingerprint USBkey and the ACC/Server. The employment of fingerprint characteristics and USBkey realize the bidirectional authentication between the user and the ACC/Server indirectly. By using password, fingerprint and USBkey formed the multifactor authentication scheme, the security is strong.

2) Anti Hacking Attacks

Analysis: In the process of the mutual authentication for User and the ACC, assuming that the attacker C monitored authentication data of User and ACC. Attacker monitored \( M_2 = E_k [E_y (n, h(ID, PW))] \) in the process (2) of authentication phase, however, as a result of lacking of private key \( x_i \) of user, it's impossible for C to decrypt arithmetic, as well get \( n \), \( h(ID, PW) \) to implement attack. In the process (5) of authentication phase, attacker C monitored \( M_5 = E_k [E_y (n, c, f_d)] | ID' \). However, the attacker C cannot achieve the purpose of hacking attacks for lacking of private key \( x_i \) of ACC.

3) Anti-Replay Attacks

Analysis: The mutual authentication relationship among user and the ACC, all of the authentication information uses the random number, such as \( (n,c) \), ensuring the freshness of information so that the attacker can't do replay attack.

4) Anti Counterfeiting Attacks

Analysis: The attacker C in the identity authentication of user and ACC, faking certification entity that is User or ACC, But the attacker C is without the private key of authentication on both sides, so C can't make fake attacks for identity authentication. When the attacker C fakes User to proceed identity authentication, because there is no User's fingerprint, C cannot use fingerprint USBkey, even cannot use fingerprint USBkey to achieve identity authentication with sever. When the attacker counterfeits the ACC, he can acquire \( M_5 = E_k [E_y (n, c, f_d)] | ID' \), but without the private key of ACC, he is unable to obtain certificate.

5) Anti Dos Attacks

Analysis: The certification request what is sent to the ACC must contain the correct \( f_d \), i.e. the serial number of user’s fingerprint characteristics. Only responding the next identity authentication, the ACC has to testify \( M_5 = E_k [E_y (n, c, f_d)] | ID' \), or will close certification sessions immediately to resist DOS attack effectively.

According to above analysis, this scheme and security contrastive analysis of literature mentioned in the introduction is shown in Table 1.

B. Performance Analysis

The results of identity authentication scheme compared with the efficiency of the scheme are presented in table 2. Using the secret algorithm SM2 to encrypt authentication messages, improves the security and efficiency of authentication protocol.
JiaoTong University. The authors would like to thank the financial support by scientific research project fund of Jiangxi province (GJJ14371) and the aids from East China Normal University. The solution design using USBkey for remote server authentication can realize remote server to the entity authentication of the user. Building digital signature is stolen easily, and can carry out USBkey to the entity authentication of users. All in a word, this scheme based on SM2 and fingerprint USBkey has high security and efficiency.

ACKNOWLEDGMENT

The author would like to thanks for the financial support by scientific research project fund of Jiangxi province (GJJ14371) and the aids from East China JiaoTong University. The authors would like to thanks the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

Zhang Lihua received the BS degree in Electrical Engineering from the East China Jiaotong University, Nanchang, China in 1994, and the MS degrees from the school of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang, China in 2003. He received the PhD degree from the Beijing University of Aeronautics and Astronautics, Beijing, China in 2011. He is currently an postdoctoral researcher in the School of Electronic Information and Electrical Engineering at Shanghai Jiaotong University, Shanghai, China, and also an associate professor in the School of Software at East China Jiaotong University, Nanchang, China. His current research interests include information security, security protocol design in wireless mesh networks and wireless sensor networks, identity authentication and mobile communication. He is a member of the IEEE and ACM.
Multiscale Recognition Algorithm for Eye Ground Texture Based on Fusion Threshold Equalization

Zhongsheng Qiu  
Naval Aeronautical and Astronautical University/Department of Traning, Yantai 264001, China  
Email: overflyq@yahoo.com

Fangfang Zhang  
Henan University of Urban Construction/Institute of Computer Science and Engineering, Pingdingshan 467036, China

Yuan Zhou  
Naval Aeronautical and Astronautical University/Department No. 7, Yantai 264001, China

Abstract—The eye ground texture is disturbed by non ideal imaging factor such as noise, it will affect the clinical diagnosis in practice, an improved multi scale retina eye ground texture recognition algorithm is proposed based on fusion area threshold. The nonlinear sampling multi-scale transform is used to analyze the geometric space coefficient of retinal vessels with multi direction and shift invariant features, the regional threshold filtering is integrated, it is used to suppress the effect of non-uniform blocks for texture recognition. The maximum likelihood local mean standard deviation analysis is used for texture parameters estimation and recognition. The noise reduced greatly, accurate identification of texture feature is obtained. Simulation results show that the algorithm can well characterize the retinal vascular texture, it has good performance in different texture feature recognition, the recognition accuracy is improved, and it has good robustness.

Index Terms—Multiscale; Non-Subsampled Contourlet Transform; Threshold Equalization; Eye Ground Texture

I. INTRODUCTION

With the rapid development of the theory of computer vision and medical image processing technology, texture feature of medical images can be used to help clinical diagnosis, the medical image processing algorithms and technology are obtained more and more attention from people. The medical images texture extraction method is becoming the trend in medical image processing. The eye ground retina blood vessel texture can predict various diseases in the human body in potential, using computer vision image processing technology, the clinical researchers and doctors can analyze the diseases and lesions trend of human body from the original multi hierarchy analysis process of texture data. The pathological phenomena and other interested regions can be analyzed in quantitative research. But the eye ground texture is often influenced by noise and other non ideal imaging factors that influence the clinical diagnosis.

Therefore, eye ground texture images processing has been the hot issue in the field of medical image processing.

Medical eye ground texture images recognition is referred as a kind of special clinical effective means to diagnosis disease. Due to noise and other non ideal imaging factors interference. The eye ground texture edge contour characteristics are basically concentrated in the high frequency image, the transform domain is susceptible to noise interference. Traditional texture noise recognition algorithm used the smoothing noise processing method, the eye ground texture information and important detail information are lost, the edge feature cannot be extracted effectively, it takes the difficulty for the subsequent image processing. In reference [4], a retinal texture recognition method is proposed based on mathematical morphology wavelet, the method adopts mathematic morphological processing method to process the texture gradient segmented image, and then the image equalization filtering is taken, the texture recognition effect is improved to a certain degree, but the texture recognition is not clear. In reference [5], eye ground texture multi scale recognition algorithm is proposed based on global threshold ridgelet transform, different threshold values are used for binary processing of eye ground texture image. The blood vessel is recognized in the binary image. For parts of blood vessel textures, the recognition effect is good, but this method needs to divide the eye ground textures into different parts, the contour information features are taken with key segmentation, this method largely improves the texture recognition features, but the computing cost is big, and the high frequency suppression effect is poor, the computational complexity is high.

In this paper, the fusion threshold equilbrium and non sub-sampled multi-scale Contourlet transform are introduced in the eye ground texture recognition. Because eye ground texture is disturbed by non ideal imaging
factor such as noise, it will affect the clinical diagnosis in practice, an improved multi scale retina eye ground texture recognition algorithm is proposed based on fusion area threshold. Simulation results show that the new method has good robustness performance, and the recognition accuracy in improved greatly.

II. BASIC KNOWLEDGE AND ALGORITHM PRINCIPLE

A. Non Sub-sampled Contourlet Transform

The non subsampled Contourlet transform a kind of special image multiscale geometric space analysis method, Contourlet transform maps the image in the multi-scale and multi direction two-dimensional space, and the multi-scale analysis and multi direction analysis are used for two relatively independent process, the image pre-processing is obtained, this process is called Pyramidal Directional Filter Bank (PDFB).

The inner product mathematical expression of subsampled Contourlet transform is shown as:

\[ s_j = \langle x, \psi_j \rangle \]  
(1)

where, \(\{\psi_j\}_{j=0}^{N-1}\) is the multi-scale pyramid directional filter banks, it is set in base space, the space coefficient can be determined by weighting computing, and \(x\) is the original image.

Tower direction filter has some limitations on the low-frequency part of the image processing, so before the tower direction conversion, the low frequency component of the filtered image is processed, and the tower direction filter interference is eliminated, the mathematical expressions after the transformation is shown as:

\[ t_j = \langle x, \phi_j \rangle \]  
(2)

where, \(\{\phi_j\}_{j=0}^{N-1}\) direction filter coefficient space, \(\mathbb{R}^N (N = 0, \cdots, 2^N - 1)\) is the base space, \(x\) is the original image input signal.

According to the formula as above, Contourlet transform mathematical expression is obtained as:

\[ c_{j,d} = \langle s_j, \phi_{j,d} \rangle = \langle x, \phi_j \rangle \]  
(3)

Furthermore, after the image process, the formula can be converted as:

\[ c_{j,d} = \langle x, \psi_{j,d} \rangle \]  
(4)

where, \(\langle \psi_{j,d}, \phi_j \rangle\) is the Contourlet transform coefficient of the original image space \(\mathbb{R}^{N \times N}\), \(x\) is the original image input signal.

In the transform domain, subsampled Contourlet transform and wavelet transform are used to suppress the noise effect, but the non subsampled Contourlet transform has not sampling step in the direction of filter decomposition process, the transform is translation invariant. The sampling process is taken in the low-pass filter bandwidth restrictions. It helps to hold the image contour features, and it can produce the geometric data of multi scale space under the multi direction analysis. Therefore, this paper uses the non subsampled Contourlet transform for image texture extraction and noise recognition.

B. High Order Cumulant Slice Texture Threshold Equalization Cancellation Algorithm

In this paper, the concept of higher order cumulant is proposed, for a single random variable \(x\), the characteristic function is:

\[ \Phi_x(\omega) = e^{\frac{1}{2} \omega^T \sigma^2} \]  
(5)

Among them, \(f(x)\) is the probability density function of random variable \(x\). According to the Gauss random variables, the second characteristic function is computed based on contour threshold equilibrium offset in the vertical texture for \(x\):

\[ \Psi_x(\omega) = \ln \Phi_x(\omega) = -\frac{1}{2} \omega^T \sigma^2 \]  
(6)

In the formula, \(\omega\) is the layered texture between layers of each order moment, \(\sigma^2\) is the two independent random accumulation value, \(\Phi_x(\omega)\) is the \(k\) order moments of random variable in the vertical texture, the \(k\) order quadrature is expressed as:

\[ m_k = E[x^k] = \int_{-\infty}^{\infty} x^k f(x) dx \]  
(7)

If \(m_k (k = 1, 2, \cdots, n)\) is survival, then the texture slices are fused with the layered texture. The random process accumulation is obtained with the characteristic function \(\Phi(\omega)\), the \(\Phi(\omega)\) can be expanded into Taylor series:

\[ \Phi(\omega) = 1 + \sum_{k=1}^{n} \frac{m_k}{k!}(j\omega)^k + O(\omega^n) \]  
(8)

where \(O(\omega^n)\) is the Taylor remainder. The for each order cumulant expression in stochastic process of texture slices is shown as:

\[ c_{1, r} = E\{x(n)\} = 0 \]
\[ c_{2, r} = E\{x(n)x(n + r)\} = r(r) \]  
(9)

By formula as above, it shows higher order cumulant is not sensitive to the Gauss noise characteristics, effects of high order cumulant is completely consistent to the noise, so as to achieve offset effect of visual jump eye ground texture, the eye ground texture of high order cumulant slice is analyzed, the eye ground texture jump offset algorithm is designed. Threshold equilibrium model is constructed, assumes that the eye ground texture state model is described as:

\[ Grass = \sum_{d=0}^{\infty} kd \times C_{grass} \times (1 - (T \cdot L)^2)^{d/2} \]  
(10)
where, $T$ is the tangent line of sight and terrain mesh, $L$ is the direction of light for detecting the $L$ is the direction of light, $H$ is a point of a user specified threshold, set a perpendicular to the slice network, high order cumulative slice texture of eye ground is fused in the texture fusion calculations algorithm, anti aliasing processing is used to construct the continuous visual feature graph model, texture slices of the multi-scale iterative equation is expressed as:

$$x_i(t) = x_i^1(t) + x_i^2(t) + x_i^3(t)$$

(11)

$$x_i^1(t) = \sum_{k=1}^{n} \rho_{i,k} x_i(t-k) - \sum_{k=1}^{n} \delta_{i,k} c_i(t-k) + c_i(t)$$

(12)

$$x_i^2(t) = \sum_{j=1}^{m} \sum_{k=1}^{n} \phi_{i,j} [w_{i,j}(t-k), \ldots, w_{i,1}(t-k)]^T$$

(13)

where, $x_i^1(t)$ is the variable position of single image data in state space $i$, $x_i^2(t)$ is the texture data generated from the spatial position, $j = 1,2,\ldots,n$, it shows the spatial proximity $I$ in the space position $j$, the multiscale reconstruction image of each pixel is set as 1, the quadrilateral visual range is formed, and the it varied with the angle of sight and slice gradually, alpha values decrease at the same time, the difference value of eye ground texture is calculated in multiscale adjacent network:

$$X_p(u) = F^0 x(u) = F^u [x(t)] = \int_{-\infty}^{\infty} K_p(t,u)x(t)dt$$

(14)

In the formula, $x(u)$ describes the multi-scale direction pixel feature sequence, $K_p(t,u)$ is the higher order cumulant feature, $c_i(k \geq 3)$, $F^u$ represents the horizontal direction relative to the rotation angle of $X$ axis.

III. Texture Multi Scale Recognition Based on the Fusion of Region Threshold

A. Algorithm Description

In the clinical diagnosis, retinal texture is restricted by many non ideal imaging conditions, the actual eye ground images is often accompanied by the interferences such as light uniform, vascular contrast distortion and contour texture fuzzy phenomenon, a certain degree of interference is produced, and even leads to a certain probability of misdiagnosis. Retinal texture of the original image data is shown in figure 1. In the transformation process of Contourlet, the non subsampled Contourlet transform abandoned the transformation process in sample link, the low frequency sub-band would not have the spectrum aliasing phenomenon. The image denoising effect is improved. The multi-scale and multi direction space geometry analysis are taken for the eye ground image.

In order to facilitate the research and analysis of the eye ground image feature recognition, the original image is taken based on figure 1. non subsampled Contourlet transform texture decomposition is taken, and the decomposition texture process is shown in figure 2. The non subsampled Contourlet transform is used for multiscale geometric analysis of the image, because the noise will make the multiscale coefficient amplified, it easily leads to loss of detail features of image edge contour edge. the soft threshold function denoising can make the image restoration is relatively smooth, but it will cause the edge details of image content loss. Threshold filtering fusion is taken for suppression of effect of maximum likelihood local mean standard deviation. The maximum likelihood estimation algorithm is used for estimation and identification of equilibrium texture parameters, it can greatest reduce the noise and the identification of texture feature is obtained.

Figure 1. Retinal texture of eye ground image

Figure 2. Texture decomposition based on non subsampled Contourlet transform

B. Regional Threshold Equilibrium

In the image threshold multiscale balance denoising, the soft threshold denoising results in a smoother image equilization, but it is easy to result in the loss of important details of image edge contour lamp. The hard threshold denoising can keep the balance edge profile lamp with more important details, but is easy to cause the image false edge phenomenon, resulting in image texture recognition error. The hard threshold equalization is applied to non subsampled Contourlet transform denoising, it can effectively suppress the image noise, due to the non subsampled Contourlet transform sampling procedure, can effectively make up for the defects of hard threshold denoising. With the adaptive threshold adjustment, it makes the image multiscale geometric analysis can maximize the retention of important details of image contour, this method can maximize the retention of important details of image contour. Thus the respective advantages of the multi-scale soft threshold and hard threshold are balanced for denoising equilibrium.
The non subsampled Contourlet transform is taken for noise image recognition and texture extraction, mathematic noise variance function with coefficients of different scales and different direction for the image is shown as:

$$\sigma_s(j,k) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n}^2(j,k)$$

(15)

Among them, \(M\) and \(N\) are the length and width of image transform domain sub-band, \(c_{m,n}(j,k)\) is the non subsampled Contourlet transform coefficients of Gauss white noise image.

The noise image is processed with non subsampled Contourlet transform, the standard deviation \(\sigma_s\) is different in different size in different direction. According to the geometrical multiscale image analysis of texture features at different scales and different directional subbands, the maximum likelihood criterion is used for estimation of noise variance in local area.

$$\sigma^2_{st}(j,k) = \max\left\{ \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n}^2(j,k) - \sigma^2_s(j,k), 0 \right\}$$

(16)

On the basis of adaptive threshold equilibrium, denoising threshold can be expressed as:

$$T = \frac{\sigma^2_{st}(j,k)}{\sigma_s(j,k)}$$

(17)

The eye ground texture is processed with non subsampled Contourlet transform, the transform domain coefficients of adjacent space has certain relatedness, multiscale geometric image noise is analyzed, abundant contour information is extracted from the different subband coefficients space, the subband has more energy, edge feature contains more obvious. The useful feature can be extracted from the region of interest or diseased part segmentation for texture recognition. In different subbands, the edge information of energy is smaller, and it contains less image edge details, the threshold equalization filtering algorithm is relatively weak, which can preserve image edge details more efficiently. Grain noise is filtered effectively. Therefore, in the process of multiscale geometric analysis of noise image, the effect of non-uniform blocks for texture recognition, and the shift invariant features are obtained. The integration of regional threshold filtering is obtained to suppress the effect of non-uniform blocks for texture recognition, and the local mean standard deviation is calculated by the maximum likelihood estimation method. The estimation and recognition of equilibrium texture parameters are obtained, it can greatest reduce noise, and the accuracy of texture feature recognition is improved.

IV. SIMULATION AND RESULT PERFORMANCE ANALYSIS

In order to verify the good performance of this algorithm, the experiments are taken. The test platform is taken based on the PC machine, CPU is Intel® Core™ i7-2600@3.40 GHz, memory is 4*4 GB DDR3@1600 9-9-9-24, the operating system is Windows7. The development tool is VS2008. The parallel processing algorithm is developed based on OpenMP 2 and MPICH NT 1.2.5. Linear algebra algorithm uses Armadillo. Figure 1 is taken as the reference image, different algorithms are taken in the simulation, the simulation results are shown in Figure 3. Wherein, (a) represents the original image, (b) is the recognition result of Wavelet algorithm, (c) is Contourlet algorithm, and (d) is NSCT algorithm.

In figure 3, several algorithms are taken for texture recognition, from the simulation results comparison, we can conclude that this method can preserve the details information in denoising, the majority of image textures have shown a clear structure, and the tiny texture information has also been better recovery.

In figure 4, the local texture screenshots contrasts results are shown, from the subjective vision, we can see that image details can be restored, and the image details results are more clear, the visual effect is better than other methods.

In order to analyze the performance of the new method, peak signal to noise ratios of image denoising are
and lesions trend of human body from the original multi
clinical researchers and doctors can analyze the diseases
computer vision image processing technology, the
various diseases in the human body in potential, using
original image can be maintained clearly.

The noise of eye ground image is filtered effectively, and the texture detail information of
is most obvious. The noise of eye ground image
region. SNR is best for new method, the algorithm of SNR effect
improvement factor are
calculated for different areas and different algorithms,
and the results are shown in Table 1. In three regions of
the eye ground image, the improvement factor are
respectively 3.32, 4.73, 5.18, the improve performance of
SNR is best for new method, the algorithm of SNR effect
is most obvious. The noise of eye ground image is
filtered effectively, and the texture detail information of
original image can be maintained clearly.

V. CONCLUSIONS

The eye ground retinal blood vessel texture can predict
various diseases in the human body in potential, using
computer vision image processing technology, the
clinical researchers and doctors can analyze the diseases
and lesions trend of human body from the original multi
hierarchy analysis process of texture data. The
pathological phenomena and other interested regions can be
analyzed in quantitative research. But the eye ground
texture is often influenced by noise and other non ideal
imaging factors that influence the clinical diagnosis. In
this paper, an improved multi scale retina eye ground
texture recognition algorithm is proposed based on fusion
area threshold. The nonlinear sampling multi-scale
transform is used to analyze the geometric space
coefficient of retinal vessels with multi direction and shift
invariant features, the regional threshold filtering is
integrated, it is used to suppress the effect of non-uniform
blocks for texture recognition. The maximum likelihood
local mean standard deviation analysis is used for texture
parameters estimation and recognition. The noise reduced
greatly, accurate identification of texture feature is
obtained. Simulation results show that the algorithm can
well characterize the retinal vascular texture, it has good
performance in different texture feature recognition, the
recognition accuracy is improved, and it has good
robustness, and it has good application value in
application.

REFERENCE

[1] Song Minghong, Yu Huafeng, Chen Haiyan. Application
of Improved Quantum Evolutionary Algorithm in
and Technology*, 2014, 30(1) pp. 170-173.

vessels in retinal images by piecewise threshold probing of
a matched filter response. *IEEE Transactions on Medical

Wavelet-based Video Distortion Metric and Applications.
*IEEE Transaction on Circuits and System for Video

PID controller. *IEEE Transaction on Industry Applications*,
2005, 42(2) pp. 537-543.

enhancement. *IEEE Proceedings on Vision, Image and

filter design and applications in denoising. *IEEE
International Conference on Image Processing, Genova,
Italy*, 2005, 749-752.

Adaptive Image Denoising Algorithm Based on Contourlet
Transform. *Infrared Technology (S1001-8891)*, 2006, 28(9)
pp. 552-556.

W., Grisan, E., Favaro, P., Ruggeri, A., Chaum, E.,
Textureless Macula Swelling Detection With Multiple
Retinal Fundus Images, *IEEE Transactions on Biomedical

[9] Li Tang, Garvin, M. K., Kyungmoo Lee, Alward, W. L. W.,
Kwon, Y., Abramoff, M. D., Robust Multiscale Stereo
Matching from Fundus Images with Radiometric
Differences. *IEEE Transactions on Pattern Analysis and

Extraction of Image Target. *Computer Simulation*, 2013;
30(1) pp. 409-413.

Zhongsheng Qiu, born in Shandong Linyi in the east of China, September 1979. He is a teacher of Naval Aeronautical and Astronautical University. In 2009 graduated from Naval Aeronautical and Astronautical University, major in computer science, received a master's degree in engineering.

Fangfang Zhang, born in Henan Pingdingshan in the middle of China, 1982. She is a teacher of Henan university of urban construction. In 2011 graduated from Wuhan University, major in computer science, received a master's degree in engineering.

Yuan Zhou, born in Anhui Huaihe in the east of China, January 1979. He is a teacher of Naval Aeronautical and Astronautical University. In 2009 graduated from Beijing University of Aeronautics and Astronautics, major in control theory and control engineering, received a master's degree in engineering.
An Improved Two-Dimensional OTSU Segmentation Method for Nvshu Character Image

Que Dong and Jiangqing Wang*
College of Computer Science, South-Central University for Nationalities, Wuhan 430074, China
*Corresponding author, Email: dqscuec@gmail.com.cn, wjqing2000@aliyun.com

Yangguang Sun
State Key Laboratory of Software Engineering, Wuhan University, College of Computer Science, South-Central University for Nationalities, Wuhan 430072, China
Email: ygsunster@aliyun.com

Abstract—For the structural characteristics of Chinese NvShu character, this paper proposes an improved two-dimensional (2-D) OTSU segmentation method based on the lateral inhibition network for Nvshu character image. A 2-D histogram with the gray level and the lateral inhibition level of the image was established and the maximum between-cluster variance was chosen as the criterion to select the optimal threshold. Experimental results show that the proposed method not only successfully reduced the effect of background noise, but also improved the accuracy of the character image segmentation, especially for NvShu character images with low contrast, uneven gray level of character strokes and uneven background.

Index Terms—Nvshu Character Image; Lateral Inhibition Network; Two-Dimensional Histogram; OTSU Method

I. INTRODUCTION

Nvshu, also called female scripts, is a mysterious symbol created by women which derived from the valley of the Xiaoshui River in Jiangyong County of Hunan Province, China. It is only used and circulated among women and is the unique female characters in the world. Nvshu is the tool of cultural communication for local ethnography, and history, etc. As an endangered language, Nvshu is in poor state of information processing in which many computer vision problems, such as stereo vision, motion estimation, image retrieval, and object recognition, can be solved better with reliable results of image segmentation [3]. Among the existing segmentation techniques, thresholding is one of the most popular approaches in terms of simplicity, robustness and accuracy. The implicit assumption is that the object and background are different in intensity values. Basically, there are two different aspects of thresholding: global thresholding and local thresholding. As the global thresholding methods are easy to be implemented, they are preferred to local thresholding methods especially in machine applications.

Many global thresholding techniques were developed over the past years. Among them, OTSU method [4] is one of the best threshold selection methods for general real world images, which selects a global threshold value by maximizing the separability of the classes in gray levels. To improve the adaptability of OTSU method, many modified OTSU methods had been proposed, such as the range-constrained OTSU method [5], the valley-emphasis OTSU method [6], and the two-stage OTSU method [7]. Generally speaking, OTSU method can obtain satisfied segmentation results in many cases. However, due to the influence of objective conditions and artificial actions, the gray level histogram may not show obvious peaks and valleys. In those cases, only using the one-dimensional histogram to determine the threshold is difficult to obtain satisfactory segmentation results. For this reason a two-dimensional (2-D) OTSU method which employs point pixel information and the local average gray level of the neighborhood pixels was presented in [8]. Although the 2-D OTSU method has stronger noise robustness, it gives rise to the exponential increment of computation time. Given this, many fast 2-D OTSU methods [9–13] had been proposed to reduce the time complexity of 2-D OTSU method. However, there is a main assumption in the above 2-D OTSU algorithms that the sum of probabilities of main-diagonal distinct in 2-D histogram is approximately one, which is contrary to realism sometimes. So many post-processing OTSU methods were presented, such as the curve 2-D OTSU method [14], the oblique 2-D OTSU method [15] and the precise 2-D OTSU method [16], which reclassify the boundary region pixels to improve the segmentation results. However, in our experiments, we found that the segmentation results often have many defects by using the traditional OTSU method, such as the fuzziness, the disconnection of character strokes and the high background noise, which are mainly resulted from the
complex textures of Nvshu materials and the high noise comes from the image acquisition equipment.

Consequently, this paper proposes an improved two-dimensional OTSU segmentation method based on the lateral inhibition network for Nvshu character image. A two-dimensional histogram with the gray level and the lateral inhibition level of the image was established and the maximum between-cluster variance was chosen as the criterion to select the optimal threshold. Experimental results show that the proposed method not only successfully reduced the effect of background noise, but also improved the accuracy of the character image segmentation, especially for NvShu character images with low contrast, uneven gray level of character strokes and uneven background.

The remainder of this paper is organized as follows. In Section II, some related works, such as the structural characteristics of Nvshu character image, the lateral inhibition network and the traditional 2-D histogram are briefly introduced. The proposed modified OTSU method is presented in detail in Section III. In Section IV, experimental results are shown. Finally, the conclusion is given in Section V.

II. RELATED WORKS

A. Structural Characteristics of the Nvshu Character Image

The original Nvshu materials are hand-written, which are usually written on manuscripts, fans, handkerchiefs and pieces of paper. In order to use the information technology to salvage and protect Nvshu, we need to scan the carriers of Nvshu and preserve them in the form of digital image. However, most of the Nvshu materials have a long history and are lack of protection, so the Nvshu character image is of low quality, which increase the difficulty of image segmentation. The structural characteristics of the Nvshu character image are listed as following: (1) the proportion of the object(Nvshu character) is relatively smaller than the background; (2) the object is composed of strange strokes with a complex background; (3) font overall tilt with the diamond structure; (4) dot strokes transformed from relatively short strokes; (5) the transformed neat lines strokes, only horizontal and vertical strokes with relatively short length could be retained, other strokes are transformed into oblique strokes with declining long lines [17]. All of these features of Nvshu character image must be taken into consideration to improve the quality of segmentation.

B. Lateral Inhibition Network

Lateral inhibition is one of the basic principles in the process of information processing of the visual neural system. It is a general mechanism in the visual and tactile sensations of living things such as limuloid and human beings. Hartline and his colleagues [18] firstly found the phenomenon of lateral inhibition in their electrophysiological experiment on limuloid eyes. If every little eye in the limuloid’s compound eyes is regarded as a receptor and all the eyes are exposed to the illumination, the experimental results demonstrated that the frequency that each eye transmits is less than what it transmits when it is under the light alone. The results of experiments also indicated that the response of a certain eye is influenced not only by the stimulation it receives but also by the inhibition affects comes from the surrounding eyes. Generally, when a neuron receives stimulation and becomes excited, it will enhance the neurons closed to it and exert an inhibition influence on the distant ones, which is expressed as “enhance center and suppress neighborhood”.

In image processing, the lateral inhibition network has the following functions [19]: 1) detecting the edges of image, highlighting the borders and enhancing the contrast; 2) high-pass filter by compressing the large input range of the network to its dynamic range; 3) compensating the fuzzy image caused by refraction system; 4) fitting the blind breaks in the image [20].

Since the lateral inhibition network was discovered by the researchers, many mathematical models had been proposed to express it. Generally, the mathematical models can be divided into cyclic models and non-cyclic models based on the inhibitory effect is from the input or the output of surrounding units. Besides, the mathematical models can be classified as subtraction models and shunt models based on the inhibitory effect is implemented by the summation or the shunt of surrounding units. In order to apply the lateral inhibition network to digital image processing, Chen and OuYang [21] proposed three 2-D lateral inhibition network models based on different classification rules.

2-D subtraction non-cyclic model:

\[ r_{p,q} = e_{p,q} + \sum_{i \in R} \sum_{j \in R} k_{p,q} e_{i,j} \]  

2-D shunt non-cyclic model:

\[ r_{p,q} = e_{p,q} \times \sum_{i \in R} \sum_{j \in R} k_{p,q} e_{i,j} \]  

2-D subtraction cyclic model:

\[ r_{p,q} = e_{p,q} + \sum_{i \in R} \sum_{j \in R} k_{p,q} r_{i,j} \]

where, \( e_{p,q} \) and \( r_{p,q} \) are the input and the output of the visual receptor, respectively. \( k_{p,q} \) is the inhibition coefficient from receptor \( (p,q) \) to receptor \( (i,j) \). \( R \) is the radius of lateral inhibition field.

C. Traditional 2-D Histogram

An image is defined as a 2-D light-intensity function, which contains \( M \times N \) pixels each with a value of brightness, i.e. gray level, from 0 to \( L-1 \), where \( L \) is the number of gray levels. The traditional 2-D histogram contains the gray level of an image and its local average gray level. The gray level of a pixel with the...
coordinate \((m,n)\) is denoted as \(f(m,n)\). Let \(g(m,n)\) be the function of the local average gray level, then
\[
g(m,n) = \left\lfloor 1/9 \sum_{i=1}^{9} \sum_{j=1}^{9} f(m+i,n+j) \right\rfloor
\] (4)
where \(\lfloor \cdot \rfloor\) represents rounding.

For an image, let \(c_{x,y}\) be the total number of the occurrence of the pair \((x,y)\), which represents pixel \((m,n)\) with \(f(m,n) = x\) and \(g(m,n) = y\), then the joint probability mass function \(p_{x,y}\) is given by
\[
p_{x,y} = \frac{c_{x,y}}{M \times N} \quad x, y = 0, 1, \ldots, L-1
\] (5)
where \(\sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p_{x,y} = 1\).

\(\{p_{x,y}\}\) is the 2-D histogram of the image. By means of the threshold vector \((s,t)\), the histogram can be divided into four quadrants

![Figure 1. Traditional 2-D histogram](image)

where quadrants 0 and 1 contain the distributions of object and background classes, whereas the quadrants 2 and 3 are the distributions of edge pixels and noise pixels. However, there is a main assumption in the above 2-D histogram that the sum of probabilities of quadrants 0 and 1 is approximately one because of that most of the pixels are located in these two areas.

### III. Modified 2-D Otsu Method

**A. Gray Level and Lateral Inhibition Level (GLLIL) 2-D Histogram**

In order to simulate the lateral inhibition network and use it to serve Nvshu character image segmentation, we design a 2-D lateral inhibition network with a radius of 3:

\[
r_{p,q} = -e_{p,q} + d \sum_{i,j\in N_8(p,q)} e_{i,j} - d \sum_{i,j\in N_{16}(p,q)} e_{i,j} - d \sum_{i,j\in N_{24}(p,q)} e_{i,j}
\] (6)

where \(k_w (w=1, 2, 3)\) is the lateral inhibition intensity coefficient and the definition of other parameters in the equation can refer to (1). Then, equation (6) can be merged as follows:

\[
r_{p,q} = -e_{p,q} + k_1 \sum_{i,j\in N_8(p,q)} e_{i,j}
\] (7)

\[
r_{p,q} = -k_2 \sum_{i,j\in N_{16}(p,q)} e_{i,j}
\] (8)

\[
r_{p,q} = -k_3 \sum_{i,j\in N_{24}(p,q)} e_{i,j}
\] (9)

and we can integrate (9) as follows:

\[
r_{p,q} = -e_{p,q} + d \sum_{i,j\in N_8(p,q)} e_{i,j} - d \sum_{i,j\in N_{16}(p,q)} e_{i,j} - d \sum_{i,j\in N_{24}(p,q)} e_{i,j}
\] (10)

where \(\bar{e}_8\), \(\bar{e}_{16}\) and \(\bar{e}_{24}\) are the average lateral inhibition value of 8 neighborhood of \(e_{p,q}\), 16 neighborhood of \(e_{p,q}\) and 24 neighborhood of \(e_{p,q}\), respectively.

By using the lateral inhibition network which is shown in (10), we can establish the GLLIL 2-D histogram. The steps can be summarized as follows:

Step 1. calculate the lateral inhibition matrix based on (10);

Step 2. filter the image by using the lateral inhibition matrix and obtain the lateral inhibition image;

Step 3. normalize the lateral inhibition image and denote it by \(h(m,n)\).
Step 4. establish the GLLIL 2-D histogram. By letting \( f(m,n) = i \) and \( h(m,n) = j \), a certain coordinate \( (i,j) \) in the GLLIL 2-D histogram is denoted, i.e., the abscissa is the gray level of a pixel and the ordinate is lateral inhibition level of a pixel, respectively. \( p_{i,j} \) represents the frequency of the occurrence of the pair \( (i,j) \). It is given by

\[
p_{i,j} = \frac{c_{i,j}}{M \times N} \quad i, j = 0,1, \cdots, L-1
\]

(11)

where \( c_{i,j} \) is the total number of the occurrence of the pair \( (i,j) \) and \( \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{i,j} = 1 \).

Fig. 2 is the comparison of the traditional 2-D histogram and the GLLIL 2-D histogram. We can see that the GLLIL 2-D histogram has a large density and the distribution of gray level is more even. Besides, comparing to the traditional 2-D histogram, the 2-D histogram proposed in this paper has more pixels at the peaks and valleys, which can enhance the contrast of image and highlight the features.

![Figure 2](image_url)

Figure 2. The comparison of traditional 2-D histogram and GLLIL 2-D histogram: (a) traditional 2-D histogram (b) GLLIL 2-D histogram

### B. 2-D OTSU Method Based on GLLIL

By obtaining the GLLIL 2-D histogram, we choose the maximum between-cluster variance as the criterion to select the optimal threshold. Suppose that all the pixels in an image can be divided into two classes \( W_0 \) and \( W_1 \) by a threshold \((s,t)\), then the probabilities of class occurrence are given by

\[
W_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{i,j} = W_1(s,t)
\]

and the corresponding class mean levels are

\[
u_0 = \left( u_{00}, u_{01} \right) = \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0} \right)
\]

(14)

\[
u_1 = \left( u_{10}, u_{11} \right) = \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_1}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_1} \right)
\]

(15)

the total mean level vector of the 2-D histogram is

\[
u_T = \left( u_{T0}, u_{T1} \right) = \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_1} \right)
\]

(16)

Based on the assumption that the occurrences of image pixels in off-diagonal quadrants of 2-D histogram can be neglected, it is easy to be verified that

\[
W_0 + W_1 \approx 1, \quad u_T = W_0 u_0 + W_1 u_1
\]

(17)

The between-cluster variance scatter matrix can be formulated as

\[
G_B = W_0 \left[ (u_0 - u_T)(u_0 - u_T)^T \right] + W_1 \left[ (u_1 - u_T)(u_1 - u_T)^T \right]
\]

(18)

The trace of the scatter matrix is expressed as

\[
trG_B = W_0 \left( u_{00} - u_T \right)^2 + W_1 \left( u_{11} - u_T \right)^2
\]

(19)

so the optimal threshold \((s^*, t^*)\) is selected by maximizing \(trG_B\)

\[
(s^*, t^*) = \arg\max_{0 \leq s \leq L-1, 0 \leq t \leq L-1} \{ trG_B(s,t) \}
\]

(20)

then convert the image to the binary one

\[
f_T(m,n) = \begin{cases} 0, & f(m,n) \leq s^* \quad \text{and} \quad h(m,n) \leq t^* \\ L-1, & f(m,n) > s^* \quad \text{and} \quad h(m,n) > t^* \end{cases}
\]

(21)

where \( f_T(m,n) \) is the binary image.

However, in the process of calculating optimal threshold, it takes too much repetitive calculation to obtain every \( W_0, u_0 \) and \( u_T \). So we use the recursive method to eliminate redundant computation and improve the computing speed of the algorithm.

Equations (14) and (15) can be rewritten as follows:

\[
u_0 = \left( u_{00}, u_{01} \right) = \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0} \right)
\]

\[
= \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0} \right)
\]

(22)

so the total mean level vector of the 2-D histogram is

\[
u_T = \left( u_{T0}, u_{T1} \right) = \left( \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_0}, \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} u_{i,j} p_{i,j}}{W_1} \right)
\]
\[ u_i = (u_{i0}, u_{i1})^T = \left( \sum_{i=0}^{m} \sum_{j=0}^{n} i p_{i,j} / W_i, \sum_{i=0}^{m} \sum_{j=0}^{n} j p_{i,j} / W_i \right)^T \]  
\[ = \left( \sum_{i=0}^{m} \sum_{j=0}^{n} i p_{i,j}, \sum_{i=0}^{m} \sum_{j=0}^{n} j p_{i,j} \right)^T / W_i \]  

Let

\[ X_0 = \sum_{i=0}^{m} \sum_{j=0}^{n} i p_{i,j}, Y_0 = \sum_{i=0}^{m} \sum_{j=0}^{n} j p_{i,j} \]  

(24)

\[ X_1 = \sum_{i=0}^{m} \sum_{j=0}^{n} i p_{i,j}, Y_1 = \sum_{i=0}^{m} \sum_{j=0}^{n} j p_{i,j} \]  

(25)

so

\[ u_0 = (u_{00}, u_{01})^T = (X_0, Y_0) / W_0 \]  

(26)

\[ u_1 = (u_{10}, u_{11})^T = (X_1, Y_1) / W_1 \]  

(27)

then establish three lookup tables

\[
\begin{align*}
P(m,n) & = \sum_{i=0}^{m} \sum_{j=0}^{n} p_{i,j} \\
X(m,n) & = \sum_{i=0}^{m} \sum_{j=0}^{n} i p_{i,j} \\
Y(m,n) & = \sum_{i=0}^{m} \sum_{j=0}^{n} j p_{i,j}
\end{align*}
\]

(28)

so

\[
\begin{align*}
W_0 & = P(s,t) \\
X_0 & = X(s,t) \\
Y_0 & = Y(s,t) \\
W_1 & = P(L-1,L-1) - P(L-1,t) - P(s,L-1) + P(s,t) \\
X_1 & = X(L-1,L-1) - X(L-1,t) - X(s,L-1) + X(s,t) \\
Y_1 & = Y(L-1,L-1) - Y(L-1,t) - Y(s,L-1) + Y(s,t)
\end{align*}
\]

(29)

By figuring out the optimal threshold using the three lookup tables in (28) and 6 formulas in (29), we can reduce the time consumption of the proposed algorithm. Three lookup tables can be established through the following three iterative formulas:

\[
P(m,n) = P(m-1,n) + P(m,n-1) - P(m-1,n-1) + p_{MIN} \quad (30)
\]

where \( P(0,0) = p_{MIN} \).

\[
X(m,n) = X(m-1,n) + X(m,n-1) - X(m-1,n-1) + m \times p_{MIN} \quad (31)
\]

where \( X(0,0) = 0 \times p_{MIN} \).

\[
Y(m,n) = Y(m-1,n) + Y(m,n-1) - Y(m-1,n-1) + n \times p_{MIN} \quad (32)
\]

where \( Y(0,0) = 0 \times p_{MIN} \).

IV. EXPERIMENTAL RESULTS

To evaluate the performances of the proposed method and demonstrate its effectiveness, the results are compared with the 2-D OTSU method and the curve 2-D OTSU method. Four NvShu character images with different typical characteristics were selected in our experiments, such as low contrast, uneven gray level of character strokes, uneven background and high noise. In the experiments, by setting \( d_w = h/2^w (w=1,2,3) \) and putting \( d_w \) into (10), we obtain the lateral inhibition matrix:

\[
\begin{bmatrix}
-0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 \\
-0.005 & -0.016 & -0.016 & -0.016 & -0.016 & -0.016 & -0.005 \\
-0.005 & -0.016 & 0.0625 & 0.0625 & 0.0625 & -0.016 & -0.005 \\
-0.005 & -0.016 & 0.0625 & 0.0625 & 0.0625 & -0.016 & -0.005 \\
-0.005 & -0.016 & 0.0625 & 0.0625 & 0.0625 & -0.016 & -0.005 \\
-0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 \\
-0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 & -0.005 \\
\end{bmatrix}
\]

Besides, all the experiments are performed on a common PC with MATLAB 2010b (Intel Core2 Duo, CPU-2.2GHZ with 2GB RAM).

A. NvShu Character Image with Low Contrast

NvShu character image with low contrast shown in Fig. 3 was chosen in our experiments. As can be seen from Fig. 3(b) and (c), the 2-D OTSU method and the curve 2-D OTSU method failed in this case. Fig. 3(d) presents the segmentation result using our method in this paper. We can see that the proposed method can get a fairly clear segmentation result which is easy to be observed by human eyes. It has been further demonstrated that our method is well adapted to the low contrast of NvShu character image.

![Figure 3](image314x261to420x346)

![Figure 3](image425x261to531x346)

B. NvShu Character Image with Uneven Gray Level of Character Strokes

In the next experiment a NvShu character image with uneven gray level of character strokes shown in Fig. 4
was chosen to demonstrate the robustness of our method. As is shown in Fig. 4(d), the proposed method has the capability to fit the strokes. However, the writer’s writing habits result in the uneven distribution of character strokes’ gray level. Since our method added the lateral inhibition information into the 2-D histogram, it effectively broadened the 2-D histogram which makes the distribution of gray level more even. As a result, our method obtains the satisfying segmented characters.

![Figure 4](image)

Figure 4. Segmentation results for a Nvshu character image with uneven gray level of character strokes: (a) original image (b) segmentation result using 2-D OTSU method (c) segmentation result using curve 2-D OTSU method (d) segmentation result using the proposed method

C. Nvshu Character Image with Uneven Background

A Nvshu character image with uneven background is shown in Fig. 5. The uneven background is resulted from the paper texture of the old Nvshu materials. As are shown in Fig. 5(b)-(d), the 2-D OTSU method and the curve 2-D OTSU method performed poorly on this image while our method successfully extracted all the Nvshu characters. This is mainly because that the proposed method utilizes the function of lateral inhibition network which can enhance the difference between the character strokes and the background. However, our method not only avoids the loss of the character image information, but also improves the accuracy and robustness of the character image segmentation. Experimental results further demonstrated its capability to overcome the effect of uneven background on character image segmentation.

![Figure 5](image)

Figure 5. Segmentation results for a Nvshu character image with uneven background: (a) original image (b) segmentation result using 2-D OTSU method (c) segmentation result using curve 2-D OTSU method (d) segmentation result using the proposed method

D. Nvshu Character Image with High Noise

In this section, we studied the anti-noise capability of the proposed method. The experiment was carried out on a Nvshu character image by adding salt & pepper noises with a density of 0.08 shown in Fig. 6(a). Figure 6(b) and (c) present the corresponding results using the 2-D OTSU method and the curve 2-D OTSU method. The segmentation result provided by our method is shown in Fig. 6(d). As can be seen from Fig. 6, the satisfying segmented characters can be obtained by using the proposed method compared with the 2-D OTSU method and the curve 2-D OTSU method. Furthermore, in the noise image, the total number of object pixels is 2259. In the segmentation results of each method, the object pixels number of the 2-D OTSU method, the curve 2-D OTSU method and the proposed method are 24632, 27576 and 1027, respectively. Misclassification rate of each method is 79.8%, 90.3% and 4.4% respectively. Our method’s misclassification rate is significantly lower than the rest two algorithms, which indicates that the proposed algorithm has the capability to reduce noise sensitivity.

![Figure 6](image)

Figure 6. Segmentation results for a Nvshu character image with high noise: (a) original image (b) segmentation result using 2-D OTSU method (c) segmentation result using curve 2-D OTSU method (d) segmentation result using the proposed method

V. CONCLUSIONS

In this paper, a improved 2-D OTSU segmentation algorithm for Nvshu character image is proposed. This novel method establishes a new 2-D histogram which contains the gray level information and the lateral inhibition information of the Nvshu character image. And then, we apply it to the 2-D OTSU segmentation algorithm. Experimental results demonstrated its
capability to reduce noise sensitivity and improve the robustness, especially for NVShu character images with low contrast, uneven gray level of character strokes and uneven background.

ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (60975021), the Natural Science Foundation of South-Central University for Nationalities (YZZ10006), the Open Foundation of Key Laboratory of Education Ministry of China for Image Processing and Intelligence Control (2009096), Academic Team of South-Central University for Nationalities (XTZ10002), the Science and Technique Research Guidance Programs of Hubei Provincial Department of Education of China (B20110802), “the Fundamental Research Funds for the Central Universities”, South-Central University for Nationalities (CZY12007, CTZ12023), and the Open Foundation of State Key Laboratory of Software Engineering (SKLSE20120934), the Natural Science Foundation of Hubei Province of China (2012FFB07404). The authors would also like to thank the anonymous reviewers for their valuable comments and constructive suggestions, which helped improve the quality of this paper.

REFERENCES


Que Dong received his B.S. degree in communication engineering from Wuhan Donghu University in 2012. Now, he is pursuing his M.S. degree in computer science and technology at South-Central University for Nationalities. His current research interests are image processing and computer vision.

Jiangqing Wang received the B.S. and M.S. degrees in Artificial Intelligence from Wuhan University, China, in 1986 and 1989, respectively; and Ph.D. degree in intelligent computation from Wuhan University, China, in 2007. She was a visiting professor of University of Wisconsin-La Crosse and Chonbuk National University. She is currently a Professor in College of Computer Science of South-Central University for Nationalities. She has published over 40 papers in international journals and conferences in the areas of artificial intelligence and intelligent computation. Her current research interests are in the areas of character recognition, intelligent computation and optimization. The research activities have been supported by the Natural Science Foundation of China, Natural Science Foundation of State Ethnic Affairs Commission and Natural Science Foundation of Hubei province.
Dr Wang has been actively involved in around 20 international conferences, serving as Session Chair and a reviewer for numerous referred journals and many international conferences.

Yangguang Sun received the M.S. degree in computational mathematics, and the Ph.D. degree in pattern recognition intelligence system, from the Huazhong University of Science and Technology, China, in 2005 and 2009, respectively. He is currently a lecturer in the College of Computer Science, South-Central University for Nationalities, China. His current research interests are image processing, computer vision, and route planning.
Improved HT Object Detection Algorithm Based on Canny Edge Operator

Tingnian He¹,², Xiaohong Li¹, and Yun Jiang¹
¹. College of Computer Science and Engineering, Northwest Normal University, Lanzhou Gansu, 730070, China
². School of Information Science and Technology, Beijing Normal University, Beijing, 100875, China

Abstract—According to the problem of low positioning accuracy of the existing line detection algorithm and multiple response problem of the single edge detection, this paper presents an improved Hough transform based on Canny operator object detection algorithm which uses the characters of the line segment object from the commonly used edge detection operator to find the best comprehensive properties of the operator. The algorithm can effectively remove the interference of noise, solved the calculation accuracy and computation speed between the optimal matching problems, and effectively solve the problem of multiple peak detection and false peak problem, improve the line segment detection robustness. Experimental results show the algorithm robustness to image noise has been improved, can get the better results.

Index Terms—Object Detection; Hough Transform; Canny Operator; Machine Vision

I. INTRODUCTION

The research object of machine vision detection has become the research focus the integration of multidisciplinary, multi field based on, by the majority of scholars [1]. In real life, people can complete object detection work easily, but the use of computers for object detection, has been plagued by researchers. The main difficulties in the detection algorithm itself is not mature enough, not enough complete; complexity of image object detection tasks [2].

Previously, some researchers have put forward some object detection algorithm, but because the computer at the time of development limit, caused the algorithm test results is not ideal; later, the rapid development of computer vision, greatly improved performance, for complex background image accuracy and the algorithm of object detection efficiency has been greatly to improve the [3]. The research object of machine vision detection has become the research focus the integration of multidisciplinary, multi field based on, by the majority of scholars. The current object detection system challenges are: changes in the object structure change the object characteristics, characteristics of the object data is external noise interference, which makes the algorithm accuracy is not high; in the complex background, object detection and tracking accuracy and detection efficiency to be modified, at the same time algorithm also requires real-time processing capacity [4].

Digital image analysis tasks is to detect object information from image sequences, and then test and estimate, finally analysis the structural parameters of objects [5], is one of the hot research field in image processing and computer vision, at present, many researchers are concerned about its application in economic and military field.

In 1962 by Paul Hough in the form of patent issued Hough transform, is mainly used to make line detection in standard linear equations of the form. In 1981, Deans improved the Hough transform, and proposed that when the standard linear equations to solve the slope under the condition of infinite parameter, the processing of the Hough transform from straight line detection extends to parametric curve [6-7].

To utilize corresponding features between dots and line segments in image space and parameter space to carry on cluster analysis, to use relation in parameter space to express corresponded dots and lines of in image space, afterwards, to detect lines by accumulator peak value in statistical parameter space [8]. Hough transformation algorithm regards pixel connection as weights, this algorithm has certain improvement for pattern noise robustness, can get some better test results. In [9] researchers have proposed to improve the standard Hough transform is a kind of parameter space to break up, according to the separated determine the method of the corresponding line, especially suitable for the intersecting line detection. Can't detect line endpoint and the length is the insufficiency of Hough transform.

The common form of line segment detection is edge detection, it is the primary step of digital picture processing [10], in order to fully keep object edge information, edge detection is able to simply processing procedure at utmost, and it is widely utilized in machine vision and its application area. Among line segment detection, it uses operator detection to decrease pattern noise and interference, it can improve line segment resolution ratio, simply parameter selection, has stronger adaptability.

RHT algorithm (Random Hough Transform) was presented by Xu Lei. This algorithm using the random sample with edge detection image selection method in order to solve the problem of previous algorithms in low efficiency and large storage space. Aiming at the shortcoming of SHT algorithm, the many to one mapping are introduced into the algorithm, changing the one to
many mapping ideas, greatly reducing the amount of calculation algorithm; Secondly, no traversal select the sample points in the space to be detected, but the local peak with detection of image to select test points, it can solve the problem of large amount of calculation, so as to improve the computational efficiency. But the random Hough transform algorithm in complex background image may still be some shortcomings, such as poor stability algorithm.

Normally, edge divides into 2 classes [11]: one is step edge which also called ramp type; the other is lien edge that also called roof type. It has below several test methods:

Gradient method [12] is an away that adopting function expression in the obvious alternation area of gray value, researchers utilizes the feature of gradient function to detect object edge.

Second order differential method. Gradient method uses first order differential operator to carry out edge detection [13], it is easy to cause missing edge detection when threshold selection is improper, so on the basis of first order differential operator, researchers evaluated second order derivative from first order derivative local maximum value, detected zero crossing point of second order differential operator, it is mainly Laplace operator and second order orientation derivative application. Although second order operator has its advantages, as it is easier to disturb by noise than first order operator, so it must choose better filter while carrying out filter processing.

Gaussian filter method. As second order operator is easier to disturb by noise, so it needs to combine better-effected filter to carry on edge detection, researchers combined Gaussian filter and Laplace edge detection organically, formed LoG algorithm [14]. First of all, researchers utilized Gaussian filter convolution to process image, not only carried out preprocessing and eliminating partial noise, but also carried on smoothly processing and leached bigger burr out, to modify image edge extension because of convolution operation, to postulate first order derivative threshold value to confirm edge range, and then to regard Laplace function as second order derivative zero crossing point to carry out calculation of partial gradient maximum value. when carries on smooth and gradient amplitude value calculation processing for images, it will adopt utilizing gradient orientation to compare edge pixel amplitude value with neighboring pixel amplitude value [15], by comparing high amplitude value with low amplitude value which set up in advance, to judge whether this pixel point is edge pixel point or not at a time.

ADM method. ADM algorithm [16] is the smaller algorithm of time and space complexity, it can adopt Gaussian filter to smooth image and eliminate noise by utilizing ADM algorithm; this algorithm has high detection rate and accurate edge position confirmation. For waiting detected image edge, it is also able to calculate edge pixel strength and direction, and get higher single responding edge.

There are some common characteristic after analyzing [17]: (1) detection error as possible as low (2) orientation is fully close (3) to restrain fake edge, decrease multiple response of single edge. Canny edge detected operator has the advantages of algorithm design simple, high efficiency, low complexity of algorithm time etc. Advantage, it has better preprocessed result for reducing image noise, strengthen image strength, it is one of the widely used detected operator at present. For Hough transformation, it has obvious advantages in edge fixed position, connectivity by adopting Canny edge detection algorithm, it is one of the ideal edge extraction algorithm, passing through conjoint analysis for line segment direction and edge pixel direction in Hough transformation, it can reduce noise influence of fake peak value in parameter space.

The implementation ideas of Canny edge detected algorithm is that to adopt Gaussian filter to process image in advance to smooth image burr first; then, to calculate image edge pixel gradient weight value and edge pixel orientation which has been preprocessed through operator; and then to find gradient partial maximum value in waiting image, set non-maximum value as zero, refine edge pixel as 1 pixel; finally, to utilize preprocessed double threshold value to analyze detection result, to eliminate the result which is greater than maximum value and less than minimum value, to keep the maximum value in double threshold value and form continuous edge. But as Canny edge detected operator does not pay more attention to outside edge information, so it causes low utilization efficiency of information in detection process [18].

The main contribution of this paper is to propose an improved Hough transform based on Canny operator object detection algorithm based on segments from the common characteristics of the object edge detection operator child to find the best overall performance of operators, and is made accumulator unit subdivision derivation. The proposed algorithm can effectively remove the interference of noise, solved the calculation accuracy and computation speed between the optimal matching problems, and effectively solve the problem of multiple peak detection and false peak problem, improve the line segment detection robustness.

II. PRINCIPLE OF CAN thy EDGE DETECTION

Description of Canny edge detected algorithm is as below:

(1) To calculate two-dimensional Gaussian function, it as formula (1) indicates:

$$G(x, y) = \frac{1}{2\pi \sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right)$$  \hspace{1cm} (1)

(2) To calculate first order derivative of predefined four direction of Gaussian $G(x, y)$ according to first step result:

$$G_x(x, y) = \frac{\partial G(x, y)}{\partial p} = p \cdot \nabla G(x, y)$$

$P$ is direction vector and shows as:
\[ p = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \]

VG(x, y) is gradient vector and shows as:

\[
\begin{pmatrix}
\frac{\partial G}{\partial x} \\
\frac{\partial G}{\partial y}
\end{pmatrix}
\]

(3) To analyze two dimension convolution sample \( \nabla G(x, y) \ast f(x, y) \) of \( \nabla G(x, y) \) under edge strength \( p \), decompose two one-dimensional filter, as formula (2), (3 indicates:

\[
G(x, y) = kx \cdot \exp \left[ -\frac{x^2}{2\sigma^2} \right] \exp \left[ -\frac{y^2}{2\sigma^2} \right] = h_1(x)h_2(y) (2)
\]

\[
G(x, y) = ky \cdot \exp \left[ -\frac{y^2}{2\sigma^2} \right] \exp \left[ -\frac{x^2}{2\sigma^2} \right] = h_1(y)h_2(x) (3)
\]

where \( k \) is constant.

\[
h_1(x) = \sqrt{k}x \cdot \exp \left[ -\frac{x^2}{2\sigma^2} \right]
\]

\[
h_2(x) = \sqrt{k}x \cdot \exp \left[ -\frac{y^2}{2\sigma^2} \right]
\]

\[
h_1(y) = \sqrt{k}y \cdot \exp \left[ -\frac{y^2}{2\sigma^2} \right]
\]

\[
h_2(y) = \sqrt{k}y \cdot \exp \left[ -\frac{x^2}{2\sigma^2} \right]
\]

And then bring these two samples to convolving with image \( f(x, y) \) respectively.

\[
A(i, j) = \sqrt{E_x(i, j) + E_y(i, j)}
\]

\[
\alpha(i, j) = \arctan \frac{E_x(i, j)}{E_y(i, j)}
\]

In the algorithm, \( A(i, j) \) represents the edge point \((i, j)\) strength of waiting image, \( \alpha(i, j) \) is the vertical edge direction of waiting image edge point. It produces roof ridge belt around maximum value of strength \( A(i, j) \) partial gradient amplitude value in waiting image, and then carries on the operation of non0maximum value restriction, sets non-maximum value of partial gradient amplitude value as zero. Non-maximum value restriction indicates as below formula (4):

\[
N(i, j) = NMS[M(i, j), \xi(i, j)]
\]

In formula(4), \( \xi(i, j) \) is restricted area, \( N(i, j) \) is the restriction result, to make sure the detection result and accuracy of detection result, set primary threshold value, compare restriction result \( N(i, j) \) with threshold value, set it as zero hen result is less than threshold value. Threshold value selection needs to refer to experience to prevent greater or less to avoid causing low contrast ratio and edge lost.

Canny operator is an operator which to adopt double threshold value method to carry on threshold process by using restriction result. First of all, to set two threshold value as \( \delta_1, \delta_2 \), image \( T[i, j] \) is the edge of low threshold value \( \delta_1 \) which got from cutting, it exist certain fake edges, but the edges are continuous; image \( T[i, j] \) is the edge of high threshold value \( \delta_1 \) which got from cutting, it exist less fake edges, but the edges are discontinuous. By adjusting and connecting continuously in detection results of high and low threshold value until gets complete edge.

From contrast test, it can be got by analysis that in the condition without noise interference, edge operator is able to detect out the result better in horizontal and water quality direction; but Canny operator and LoG operator are the best in smoothly eliminating noise; in accurate positioning, Sobel operator and LoG operator are the least; in missing edges, Sobel and Prewitt operator result is better, but it’s easy to occur fake edges.

### III. PROPOSED IMPROVED HT ALGORITHM BASED ON CANNY EDGE OPERATOR

#### A. Hough Transform Algorithm Analysis

It was proposed an improved algorithm in this chapter, to mix Hough transformation with Canny operator organically, and proposes to preprocess image by adopting Canny operator, and then utilize least square method to fit line segment histogram, finally, to use Hough transformation to detect line segment peak value of parameter space to accurate idea of location. According to HT, it can be known that there exists any point in digital image, if all line segments which pass this point, it could be expressed in equation \( y = ax + b \), where a represents line segment rake ratio which crosses this point, \( b \) represents line segment nodal increment which crosses this point, then we rewrite former equation into \( b = -ax + y \), and shine upon point \((x, y)\) line equation which in ab flat surface to another point \((x, y)\) in parameter space flat surface, two lines constitute a line, and suppose point \((a', b')\) is the crossing point of related lines, where \( a' \) represents rake ratio which passing this crossing point, \( b' \) represents nodal increment of line segment which passing this crossing point and point \((x, y)\) in xy flat surface.

At present, the most used Hough transformation algorithm divide in two type, one is that each sub aggregate which formed by \( R \) numbers characteristic points uses the format of several-for-one to shine upon HT to mix PHT algorithm in parameter space, as this algorithm sets two parameters \( \rho, \theta \) in parameter space, and does quantization processing for parameter, so adopting two-dimensional array notation to detect maximum point in parameter space while describing parameter space, this method is able to detect
characteristic points of line segment totally, but also there are still some more noise to interfere detection result. Another is CGHT algorithm, its performance in time complexity and space complexity are better, parameter space adopts a parameter $\theta$ to better detect out line segment of horizontal direction, vertical direction or opposite angles $(\pm \pi/4)$, as line segment pixel almost concentrate in one line, but when line segment direction is changeable, the closer the distance with detection area of line segment direction, the more noise detection eliminate; the larger the distance with detection area of line segment direction, the more noise are detected, to influence detection result, to decrease detection accuracy. These two situations have certain limitations in detection area.

B. The Process of Improved Algorithm

Improved algorithm procedure is as below:

Parameter value $\rho, \theta$ which all pixel points in one line segment correspond in image space, to let $xy$ image space corresponds with $\rho \theta$ accumulator space, that is $\rho = x \cos \theta + y \sin \theta$.

To define rank $n$, which is action of waiting detected image:

To calculate parameter extreme value:

$$
\rho_{\max} = \sqrt{n_{x}^{2} + n_{y}^{2}} \\
\rho_{\min} = -n_{x}
$$

$$
d_{\rho} = 1, d_{\theta} = \arctan \left( \frac{1}{\max(n_{x}, n_{y})} \right)
$$

To calculate dimensional formula:

$$
n_{\rho} = 2^{(\log_{2} \rho_{\max} - \rho_{\min})} \\
n_{\theta} = 2^{(\log_{2} \max(\theta) - \min(\theta))}
$$

Abscissa axis of accumulator space is $i$, accumulator unit $(\theta_{i}, \rho_{j})$ if accumulated value is $k$; it means this unit has $k$ numbers sine curve.

Derivative of sine curve:

$$
\frac{d \rho}{d \theta} = \sqrt{x_{i}^{2} + y_{i}^{2}} \cos(\theta_{i} + \phi_{i})
$$

Rake ratio of sine curve:

$$
t_{i} = \frac{d \rho}{d \theta} = \sqrt{x_{i}^{2} + y_{i}^{2}} \cdot \cos(\theta_{i} + \phi_{i}) \cdot d_{\theta}
$$

Average value of normal angle:

$$
\psi = \frac{1}{k} \sum_{i=1}^{k} \psi_{i}
$$

Angle:

$$
\psi_{i} = \arctan \left( -\frac{1}{\sqrt{x_{i}^{2} + y_{i}^{2}} \cdot \cos(\theta_{i} + \phi_{i})} \right)
$$

Used exponential function $f(x) = e^{-a x + b}$ to transform it into $g(x) = \ln(f(x)) = b - ax$.

Each seed point $(x_{i}, y_{i})$ in characteristic point assemblage calculates and modifies accumulator value, and compares with threshold value $T$ according to $\theta_{i}$ which in accumulated matrix $H(\theta)$, it is able to calculate $\rho_{i} = x_{i} \cos \theta_{i} + y_{i} \sin \theta_{i}$ on the basis of line segment parameter, to use this to judge if there exists definite line segment of $(x_{i}, y_{i})$ and $\theta_{i}$.

To keep line segment parameter $(\rho_{i}, \theta_{i})$ in new array $H^{*}(\rho, \theta)$, and to delete characteristic points of line segment which has detected out from dot bank, it’s able to reduce repetitive operations, then to let additive matrix go on initial operation, make accumulator space $H(\theta)$ get recycle use, decrease algorithm space complexity by that.

To execute step 1 to step 7 cyclically, judge all seed point once in characteristic point assemblage until the characteristic point exists which do not satisfy terms, the final detection result keeps in matrix $H^{*}(\theta)$, then the algorithm finished.

It is discovered by this algorithm analysis that it needs to carry on preprocessing operation first, such as color threshold value division and outline extraction, and then applies for HT transformation. To improve detected waiting image quality and reduce noise, also can use space filter to operate, but space filter can’t clear efficiently or edge noise, as this kind of noise is nothing wrong with image quality, so algorithm should as certain Robustness which overweight in edge noise.

![Figure 1. The flow chart of lane detection based on proposed algorithm](image)

C. The Lane Line Detection Steps

Detection step of traffic lane which based on HT algorithm

It includes 4 steps about traffic lane detection which based on HT algorithm, as figure1 indicates.

First, it’s in interest area, to extract Canny edge for waiting detected image to get edge image and edge direction image in edge pixel.
Then to combine edge direction image to let edge pixel of waiting detected image carry on Hough transformation, make it transform into Hough parameter space, let line characteristic parameter carry on cluster operation in Hough parameter space.

Afterwards, to search line segment characteristics which are greater than certain threshold value in Hough parameter space in interest area, to carry detecting and orienting primarily.

Finally, to analyze the result of traffic lane detection and get final statistical information.

Line analytic equation expresses by polar form, parameters are respective drift angle \( \theta \) and centerline spacing \( \rho \), and parameter space is two-dimensional limit parameter space, as formula (10) indicates:

\[
\rho = x \cos \theta + y \sin \theta
\]  

(10)

In the procedure of application, to be convenient for specific calculation, it needs to do discretization for parameter space. The common method is to change parameter space \((\rho, \theta)\) into \(N_\rho \times N_\theta\) two-dimensional array, where vertical dimension \( \rho \) of line segment and original point \( o \) expresses in abscissa, angle \( \theta \) between line segment and \( x \) axis expresses in ordinate, here exists limitation setting method of reasonable two-dimensional array, as formula (11) indicates:

\[
\begin{aligned}
N_\rho &= \left\lfloor \frac{\sqrt{X_n^2 + Y_n^2}}{\Delta \rho} \right\rfloor + 1 \quad 0 \leq \rho \leq \frac{\sqrt{X_n^2 + Y_n^2}}{2} \\
N_\theta &= \frac{2\pi}{\Delta \theta} \quad 0 \leq \theta \leq 2\pi
\end{aligned}
\]  

(11)

In above formula, \( \sqrt{X_n^2 + Y_n^2} \) represents diagonal length, \( \Delta \rho \) represents vertical distance which each pixel point from own line segment to original point \( o \) in abscissa direction, \( \Delta \theta \) represents angle difference which each pixel point from own line in ordinate direction.

Formula (12) gave computational formula which based on improved HT: The edge direction of edge point and HT angle relationship:

\[
H(\rho, \theta, x, y, \varphi) = \begin{cases} 
\text{abs} (\cos (\theta - \varphi)) \\
0
\end{cases}
\]  

(12)

Therefore, algorithm steps of Canny edge and HT in interest area:

1. To calculate traffic edge image and edge image angle, as formula (13) indicates:

\[
\begin{aligned}
I_{\text{edge}}(x, y) &= \sqrt{P(x, y)^2 + Q(x, y)^2} \\
I_{\text{angle}}(x, y) &= \arctan (Q(x, y) / P(x, y))
\end{aligned}
\]  

(13)

2. To calculate traffic lane edge image in interest area, as formula (14) indicates:

\[
I_{\text{edge}}' = I_{\text{edge}} * I_{\text{mark}}
\]  

(14)

(3) To carry on Hough transformation in traffic lane edge of interest area, as formula (15) indicates:

\[
\begin{aligned}
Hough(x \cos \theta + y \sin \theta, n) &= Hough([x \cos \theta + y \sin \theta], n) + \text{abs} (\cos (\theta - I_{\text{angle}}(x, y)))
\end{aligned}
\]  

(15)

By taking pictures and general knowledge, it can be seen that linearity in traffic lane is strong, these tangential paths in Hough space is concentrated in partial peak value normally, so it is able to seek partial peak value in Hough space to check traffic lane edge.

In Hough space, maybe there are some fake peak value which occurred because of distribution position, if only because of single threshold value to extract out this set peak value, maybe decrease identity productivity, therefore, it can set partial peak value to go on restriction while carrying on first Hough space filtering; When carries out second filtering to extract characteristics, to restrict by setting closest distance of characteristic parameter space.

In this chapter, it was proposed a new improved HT algorithm for HT detection algorithm shortage, such as bad automaticity in detection procedure which needs human intervention, to miss line and occur fake line segment while detecting, weak for noise Robustness etc., and problems. According to line segment and noise feature, to utilize Canny operator to strengthen expected peak value for distribution difference of peak value in accumulator space, adopting index curve to fit further to reduce noise to get better Robustness threshold value. The result shows this algorithm with strong automaticity, does not need human intervention, low frequency in missing line and occurring fake line segment, strong Robustness.

IV. RESULTS OF SIMULATION AND PERFORMANCE EVALUATION

Intelligent transport navigational system, traffic lane detection and tracking is a necessary function, it can be as a reference to prevent car deviating lane guidance, also the precondition of obstacle detection. Image preprocessing mainly has: image filtering, image strength, image binary. The experimental data coms rom data which driving recorder photographed, for simulated noise, to contrast noise removal algorithm in MATLAB 7.0 toolbox and of improved algorithm, it can get improved HT algorithm which is able to detect line segment, fast arithmetic speed, better experimental result.

A. Image Preprocessing

If waiting detected image does not carry on image preprocessing, there should exist some noise which influence image more or less, it will increase the difficulties of image detection and analysis, influence image characteristics and edge presence. So it must carry on filter processing for noise, especially isolated point noise and edge shaking noise, it will cause much influence while detecting. As image signal and noise is blend crossover in waiting detected image, if carries on smoothly processing, it will blur waiting detected image...
edge, and then decrease its quality. So for noise processing, normally, it needs to have a suitable algorithm which not only can keep waiting detected image details, but also reduce noise, and also need to give consideration to algorithm efficiency while real time detection.

Image filtering algorithm can be divided into: filtering algorithm of frequency filtering image deals with orthogonal change, needs more calculation amount, can’t satisfy real time detection speed; filtering algorithm of space filtering image is an algorithm that to let image processing algorithm use gray value, to adopt smooth image edge and mix math morphology, normally has neighborhood average method and median value filtering method. The purpose of image preprocessing is to fasten latter detected instantaneity and improve latter detected accuracy, median filtering not only can protect waiting detected image edge clarity, but also can eliminate noise influence, so it is able to adopt fast median filtering to process waiting detected image. To analyze fast median filtering, first is to choose a suitable window, this sliding window has odd numbers pixel, from the left top corner of waiting detected image, from left to right, top to bottom to slide window, arrange all pixel gray value in slipped window in small and big order, choose medium value as gray value of window center pixel point, output this value, as formula (16) indicates.

\[ f(x,y) = \text{median}[S_{f(x,y)}] \]  
(16)

In formula, \( S_{f(x,y)} \) is the area of present \((x,y)\).

Fast median filtering algorithm is as below:
To build a slide window which size is \((2N+1)*(2N+1)\), from the left top point of waiting detected image begins, along image rom left to right, top to bottom to slide window, sliding step size is 2.
When window slides one step according to certain step size, it will be able to get element rank which has \((2N+1)*(2N+1)\) numbers, then arrange all elements in slide window, potentially confirm middle pixel which mark as pixel \([(2N+1)*(2N+1)+1]/2\);
To use element \((2N+1)\) which is outside of slide window \((2N+1)*(2N+1)\), adjust and compare with confirmed median element, then can get median needle of present slide window;
To improve efficiency, when slide window moves once, not completely restart to calculate median value, here we adopt to eliminate the first element in slide window, keep the spare ranked elements, then add the new ranked element into slide window, get a modified slide window, repeat step 2,3 to get next median value;
According the sliding order of waiting detected image which rom left to right, top to bottom, to carry on step1 to step 4 operations for the whole image.
Fast median filtering algorithm evaluated pixel median value of two slide windows in an order procedure; it has a great improvement after comparing detection efficiency and speed of median filtering method, the result of contrast as figure 2 shows.

Although median value has some influence for waved signal, it do not damage step function signal and ramp signal noise, only weaken top peak value signal; but median value maybe damage some details of waiting detected image, as when it restrains random dot noise,
also will causes impulse signal damage which is less than window $1/2$, the window size and signal damage is direct proportional, so it is very important in practical application to choose a suitable size window.

It can be used Canny operator to enhance the difference between expected peak value and noise, make waiting detected image to choose threshold value in a bigger range, and result is that peak value is available. To improve algorithm general performance, it can design a general algorithm which chooses different threshold value in different situation. Figure 3(a) shows accumulated value histogram which an image passes traditional Hough transformation, figure 3(b) is histogram after figure 3(a) passed Canny operator and preprocessed image, to contrast difference between two images, it can be sure that peak value of the second image is more effective.

It can be used Canny operator to enhance the difference between expected peak value and noise, make waiting detected image to choose threshold value in a bigger range, and result is that peak value is available. To improve algorithm general performance, it can design a general algorithm which chooses different threshold value in different situation. Figure 3(a) shows accumulated value histogram which an image passes traditional Hough transformation, figure 3(b) is histogram after figure 3(a) passed Canny operator and preprocessed image, to contrast difference between two images, it can be sure that peak value of the second image is more effective.

### B. Value Processing

Edge information is the image important feature of waiting detected image, it occupies prominent position in image analysis and machine vision, it can reflect sudden changes of partial gray value of waiting detected image, color sudden changes and other features, it can distinguish different areas of waiting detected image.

Frist of all, edge extraction operation need to detect out discontinuity of image partial characteristics, then connects discontinuous edge pixel into edge, image edge characteristic is that pixel changes moderately along edge direction, vertical pixel changes sharply along edge direction. So we can utilize math algorithm to extract edge.

Relative to lane line detection, image binaryzation is key link. Figure 4 is the image impression drawing after completing image median filtering and binaryzation. Where figure 4(a1) is straight artwork, figure 4(a2) is straight preprocessing picture; figure 4(b1) is left image, figure 4(b2) is left preprocessing image, and figure 4(c2) is right preprocessing image.

After preprocess image, it is able to detect road traffic line, in normal situation, traffic lane reticule has below characteristics: 1. traffic lane is drawn by strong light-reflecting plastic; 2. Traffic lane is paralleled, certain width; 3. Traffic lane is consist of continuous or discontinuous tangential path; 4. Traffic lane edge is a line; bend is in low cavity on highway, it can process traffic lane reticule which in 100-meter car range into a line.

### C. Experimental Result and Analysis

To utilize tachygraphy as means to get vehicle information, carry on experiment for 2200 frame traffic lane picture in for different situation, background, run algorithm in the environment of VC++6.0 and take 200 seconds. It was found that when pavement condition is better, background complexity is low, the detection accuracy is high, where regard detection rate of straight lane is high, up to 91.2%, bend detection rate is up to
90.1%; When pavement condition is bad, but when interference is low, detection rate is relative low, road detection rate is 88.7%; While pavement condition is bad, background is complex, a lot of interference, detection rate is 83.2%. By analyzing, the detection result of traffic lane has certain relationship with pavement condition, sunshine condition, background complexity etc. conditions; it’s easy to be influenced.

As figure 5 shows, to test match ratio between traffic lane model and real traffic lane. If certain information is greater than threshold value, it will mark the point as eligible candidate point; if certain information is less than threshold value, it will mark the point as unqualified candidate point and retake sample. Candidate point in high matching rate will be the fitting result of last traffic lane, where candidate points can be set by experience.

V. CONCLUSION
In this chapter, it was studied object detection which based on Hough transformation, and proposed a line segment detection algorithm of improved Hough transformation on the basis of Canny edge operator. For line segment object characteristics and probationer detected operator, it carries out analysis, comparison and constitution, after studying and comparing the common used edge detected operator, to look for an operator which is with best overall performance as experimental operator in of this chapter from frequently-used edge detected operator. To carry on analysis for advantage and shortage of basic Hough transformed detected algorithm, make improvement for detection algorithm, propose line segment detected algorithm of Hough transformation on the basis of Canny edge, and do formula derivation for accumulator unit subdivision. This algorithm is able to efficiently balance time complexity and accuracy, decrease noise interference at utmost; for multiple peak value and fake peak value problem of detection result, it has been solved in a certain extent.

REFERENCES
Image Signal Enhancement based on Fractional Differential Technologies

Zuo Lijuan, Bai Cuixia, and Yang Yunfan
Polytechnic College of Hebei University of Science and Technology, Hebei Shijiazhuang, China
Email: sallyzuo2005@126.com

Zhuo Fengqin
Shijiazhuang City Economic Vocational College, Hebei Shijiazhuang, China

Abstract—To describe the ability of image fractional differential in texture detail enhancement and to study the lateral inhibition principle, multiformal masks for digital image fractional differential and their operation rules are discussed. The theoretical basis of fractional differential in modulation and demodulation is also analyzed. Through the derivation on the relation between fractional differential and signal time-frequency analysis, we can acquire the separability of two-dimensional fractional differential under certain conditions. Mach’s phenomenon generated in image texture detail is studied by us from two aspects: optic nerve model and signal processing, to propose novel masks. Then its operating rules for digital image processing based on fractional differential are proposed. The experiments show that our scheme can effectively reserve better information of edge texture detail during the denoising process, especially for weak texture and gray with little change. The total enhancing effect is obviously superior to integer order differential operator.

Index Terms—Fractional Differential; Receptive Field; Mask; Denosing; Texture

1. INTRODUCTION

Digital image acquisition and transmission will be interfered by the noise in different degrees. It will not only seriously influence the visual effect of image but also bring about difficulties on high-level processing, such as division, recognition, understanding, etc. Therefore, image denoising research is one of important contents in image processing. So far, domestic and overseas scholars have proposed several classical image denoising algorithms like mean filter, order statistics filter, low-pass filter, wiener filter and other filters which derive from these methods, including fuzzy filter, self-adaptive mean filter, filter based on edge features, etc [1-5]. Above filtering methods can remove noise from images to some extent but these methods will lose detailed information of image such as edge and texture, which will result in image fuzziness. The phenomenon lies that these image denoising algorithms are directly or indirectly applying integral order differentiation to construct denoising model. This kind of model will seriously eliminate high frequency information and it will lead that the images after denoising will lose amounts of detailed information including edge and texture, etc.

Fractional calculus is an important branch in mathematical analysis. It has been applied to signal analysis and processing, especially for digital image processing [6-8]. References [9-10] mainly use fractional differential to improve high frequency component and nonlinearly reserve special quality of signals. Thus, fractional order differentiation theory is effective in image enhancement [11]. Research shows image signal has high self-similarity. In digital image, gray value between pixels is highly correlated. The processing results of one-dimensional signals by fractional order differentiation not only enhance high-frequency, but also reserve the low-frequency and DC information of signal. So we can naturally predict that fractional order differentiation of two-dimension image signal can enhance high-frequency edge information with relatively large gray value change. Besides, it can reserve its low-frequency and nonlinearly enhance detailed information of texture without large change in its smoothing area as much as possible.

Therefore, we attempt to apply fractional order differentiation theory to image denoising and proposing the image denoising algorithm. This method breaks away traditional image denoising method based on integral integration order. Iterative method is applied to control image de-noise effect to keep detailed information of image, by setting a smaller fractional integration order to construct corresponding image de-noise model, such as edge and texture during image denoising. This paper deduces and analyzes the relationship between fractional order differentiation and other classic time-frequency researches, so as to analyze the separability of two-dimensional fractional order differentiation under certain conditions. On this basis, theoretical principles of fractional order differentiation for signal modulation and demodulation are discussed. The physical significance of fractional order differentiation in dynamic system is also proposed. Then we study the receptive field model of image fractional order differentiation. In addition, we analyze special Mach’s phenomenon of image texture detailed information from two perspectives: visual neural model and signal processing. Finally, various fractional
order differentiations and its operation rules in digital image processing are discussed. The experimental results show that the effect of image fractional order differentiation operation is obviously superior to image integral order differentiation operation when extracting and enhancing texture detailed information in smooth region. It is very outstanding for image processing with rich texture information. We have successfully implement some image fractional order differentiations on some actual cases. By comparison and analysis on the experimental results, the performance of improved scheme is fully certified.

II. RELATED WORKS

A. Relation Derivation between Fractional Differential and Time-frequency Analysis

First-order Fourier change [12] is used to describe Wigner distribution of $s(t)$ which is rotated counterclockwise $\alpha = \pi/2$ from time axis $t$ to frequency axis $\omega$, that is, it is output of a function on $\omega$; correspondingly, p-order Fourier change is the formation to describe Wigner distribution of $s(t)$ which is rotated counterclockwise $\alpha = p\pi/2$ from time axis $t$ to frequency axis $\omega$.

When the order of fractional signal Fourier change $S_p^\alpha(u)$ $p \rightarrow 0$, Wigner distribution of $s(t)$ is rotated counterclockwise $\alpha = p\pi/2^{p \rightarrow 0}$ and we acquire the factional Fourier change. If $p \rightarrow 0$, Wigner distribution of $s(t)$ nearly has rotation. At this moment, the factional Fourier change will lose the analysis ability on signal change in detail. It is consistent with the physical meaning of signal fractional differential. It is the qualitative interpretation for formula 1. On the one hand, when order of Fourier change $p \rightarrow 0$, the analysis of factional Fourier change on signal time-frequency will degenerates to that on time only; On the other hand, p-order fractional differential $s^{(\alpha)}(t)$ of signal $s(t)$ can be compensated by $p \rightarrow 0$ factional Fourier change $S_p^\alpha(\mu)$ at time $\tau = t \sec(\alpha)$. Therefore, $s^{(\alpha)}(t)$ owns strong ability to analyze the time detail on signals.

$$s^{(\alpha)}(t) = \lim_{p \rightarrow 0} \Delta(\nu, p, \tau) S_p^\alpha(\mu)$$

(1)

B. Signal Modulation and Demodulation

The typical unstable signal for time-frequency aggregation evaluation is Linear Frequency-modulated signal (LFP). The Wigner-Ville distribution of LFP signal with amplitude 1 is expressed as

$$WV_{LFP}(t, \omega) = 2\pi [\omega - (\omega_0 + at)]$$

(2)

It can be seen that Wigner-Ville distribution with unit component is impulse line spectrum which is distributed along the line $\omega = \omega_0 + at$. Its amplitude value all appear on the line representing the rate of instantaneous frequency change. As is well known, the Fourier change of p-order fractional order for signal $s(t)$ is acquired by distributed rotation $\alpha = p\pi/2$ of Wigner-Ville. From formula 1 we can see, p-order fractional order of $s(t)$ is approaching the normal weight distribution of Wigner-Ville. So it provides the theoretical basis for fractional differential to be applied on signal modulation and demodulation. Fractional differential satisfies the inhibition principle so it can be implemented in two-dimension image processing. Theories and related experiments have demonstrated fractional differential can extract the detail information of textures, whose effect is prominent especially for richly textured images [13,14].

C. Mathematical Model of Lateral Inhibition and Edge Extraction

The moment of receptive field stimulation is very sensitive to changes on time. When providing and withdrawing the light, there is big up and down in firing frequency. Receptive field stimulation is particularly sensitive to chiaroscuro edge, which shows smooth stimulus period. This special sensitive character of ganglion cells to space chiaroscuro edge can be used to explain the Mach’s phenomenon, as is shown in figure 1.

![Figure 1. Physiological basis of Mach’s phenomenon](image)

At location $a$ in figure 1, $\theta = \theta_0$. Let its corresponding sensitivity of receptive field $h = (\theta - \theta_0)$. The physical light intensity of its adjacent area is $s(\theta_0)$. So the subjective perception light intensity of ganglion cells for this receptive field is

$$y(\theta_0) = \int h(\theta) d\theta$$

(3)

Since the receptive field of human is the type of centre-excitement when $0 < S_h < 1$, we get

$$y(\theta_0) < s(\theta_0)$$

(4)

Similarly, at location $d$, its corresponding subjective perception light intensity is
\[
y(\theta) = \int_{\theta_e}^{\theta} s(\theta) h(t - \theta) d\theta = s(\theta_e) S_h - s(\theta) (5)
\]

At location \( b \), \( \theta = \theta_e \). When the width of receptive field is less than the area for physical light changing, we have

\[
y(\theta_e) = \int_{\theta_e}^{\theta} s(\theta) h(t - \theta) d\theta = s(\theta_e) S_h \quad (6)
\]

Similarly, at location \( c \), \( \theta = \theta_i \)

\[
y(\theta_i) = s(\theta_i) S_h - K \gamma_k
\]

\[
y(\theta_i) - \delta_i < y(\theta_e) (7)
\]

After variable substitution,

\[
y(\theta) = \int_{\theta}^{\theta_e} s(\theta) h(t - \theta) d\theta = s(\theta)^* h(\theta) (8)
\]

Without loss of generality, we get

\[
y(t) = \int_{t}^{\theta} s(\tau) h(t - \tau) d\tau = s(t)^* h(t) (9)
\]

Then the viewpoint of signal processing can be used to analyze the Mach’s problem.

III. Digital Image Signal Enhancement Method Based on Fractional Differential

A. Theoretical Analysis of Numerical Implementation

Without loss of generality, we take \( v = 0.5 \) order differential of the sawtooth wave and the square wave signal as examples. Their experiment results are shown in figure 2. By analysis of these two figures, though there are different definitions of fractional differential in mathematical function, all of them should satisfy the following characters:

The value of fractional differential in flat sections corresponds to the maximum value of singular transition which approaches zero gradually. While any integer order differential in flat sections must be zero.

The fractional differential at gray ladder or starting point of the slope is not zero, which strengthens the high-frequency information.

The value of fractional differential along the slopes is not zero or constant, while that of integer differential is a constant.

Thus, the fractional differential can reserve part of low-frequency contour information, when strengthening the image signal and detailed information of the texture. Compared to integer differential, fractional differential has favor in extracting edges and keeping contours. The richly textured signals contain highly self-similar fractal information. So during the process of strengthening texture details, fractional differential has its advantage in theory. In addition, since the fractional differential at flat gray area is not necessarily zero. After fractional differential the images will get relatively thicker edge transition than integer differential.

B. Antagonism Analysis and Texture Details Extraction

According to the character of Fourier change, its one-order derivative \( s'(t) \) has equivalent form

\[
Ds(t) \Leftrightarrow (D_s)(\omega) = (i \omega)^n \hat{s}(\omega) = \hat{d}(\omega) \hat{s}(\omega) \quad n \in R^+ (10)
\]

In complex region its index form is:

\[
\hat{d}(\omega) = (i \omega)^n \hat{s}(\omega) = \hat{a}_n \exp(i \hat{\theta}_n(\omega)) = \hat{a}_n \hat{\phi}_n(\omega)
\]

\[
\hat{\theta}_n(\omega) = \left| \omega \right|^n, \hat{\phi}_n(\omega) = \frac{n \pi}{2} \text{sgn}(\omega) (11)
\]

In formula 14, we can think the physical meaning of signal fractional differential is generalized phase and amplitude modulation. The amplitude changes with the exponent of frequency ingredient and its phase is generalized Hilbert change of frequency. So the problem is whether we can construct a two-dimension image operator based on fractional differential, to strengthen the texture detailed information and find the highly self-similar fractal information.

© 2014 ACADEMY PUBLISHER
Assume the general antagonist features is \( h(t) \). The light intensity of subjective intuition \( y(t) \) is the convolution of physical light intensity \( s(t) \) and \( h(t) \).

\[
y(t) = \int \! s(\tau) h(t-\tau) d\tau = s(t) h(t)
\]  

(12)

If fractional differential is used for the image signal, the impact response of filter \( d_n(t) \) equals to its corresponding general antagonist feature of the bionic receptive field, \( h(t) = d_n(t) \). When \( 0 < n < 1 \), at band \( \omega > 1 \), the fractional differential has less function in amplitude strengthening for high-frequency components than one-order differential. But in the frequency band which approaches zero frequency like \( 0 < \omega < 1 \) or lower frequency band, the signal fractional differential will not generate large linear attenuation. It will get nonlinear attenuation and the degree is not too big. Therefore, for fractional differential with \( 0 < n < 1 \), the central mechanism \( h_{\Omega}^{m0}(t) \) of bionic Rodieck receptive field should have less amplitude and flatter results than those of integer differential. Simultaneously, the peripheral mechanism \( h_{\Omega}^{m0}(t) \) of image fractional differential should have less amplitude in general antagonist features. It has longer tail and weaker compact support. Corresponding visual width of receptive field \( \Theta_{\Omega}^{m0} \) is bigger than that of integer differential \( \Theta_{\Omega}^{m1} \).

C. Approximate Structure

Generally, we apply the differential expression of unary function to two-dimension. The x-direction and y-direction of two-dimension fractional differential [12] are defined as:

\[
\frac{\partial^n f(x,y)}{\partial x^n} \approx f(x,y)+(-a)f(x-1,y)+\frac{(-a)(-a+1)}{2}f(x-2,y)+...+\frac{\Gamma(-a+1)}{n!\Gamma(-a+1+n)}f(x-n,y)
\]  

(13)

\[
\frac{\partial^n f(x,y)}{\partial y^n} \approx f(x,y)+(-a)f(x,y-1)+\frac{(-a)(-a+1)}{2}f(x,y-2)+...+\frac{\Gamma(-a+1)}{n!\Gamma(-a+1+n)}f(x,y-n)
\]  

(14)

It can be seen that for the approximation of partial differential for function \( f(x,y) \), the parameter value of each corresponding sum item are the same. Thus, they can sequentially replace zero value in \( n \times n \) all-zero matrix. According to above two formulas, we can construct \( 3 \times 3 \) fractional differential mask operator. It is used to work on the negative direction of x-axis and the positive direction of y-axis in two-dimension mask. The mask operators need to be added to left and right diagonal direction, to enhance the anti-rotation performance. In addition, there are 4 mask operators in vertical direction of left diagonal and right diagonal. So there are totally 8 mask operators in all directions. They are added to get the final mask as figure 3(e).

![Fractional differential mask](image-url)

**Figure 3.** Fractional differential mask
The digital image processing is based on the process of discrete pixels. The computing rule of fractional differential uses the scheme of fractional differential convolution to make spatial filtering. Its mechanism is that the digital images waiting for process will be covered by point by point. The numerical computing rule of digital gray image is described in the following figure and it is similar to the rules in the negative direction on x-axis.

(a) There is only constant coefficient of fractional differential mask in 8 directions. Its coordinate \((x, y)\) and pixel \(s(x, y)\) waiting for fractional differential must be kept coincided.

(b) Multiply the parameter value of fractional differential mask by gray value of their corresponding pixels. Then all the products are added to acquire the weighted sum.

(c) Since the gray value of digital image is between 0 and 255, we need limit the amplitude of the weighted sum. When its value is less than 0 the amplitude is zero; when it is more than 255, the amplitude is 255.

\[
\begin{align*}
S(x, y) &= s(x, y) \\
S(x-1, y) &= s(x-1, y) \\
S(x-1, y-1) &= s(x-1, y-1) \\
S(x-1, y+1) &= s(x-1, y+1) \\
S(x, y-1) &= s(x, y-1) \\
S(x, y+1) &= s(x, y+1) \\
S(x+1, y) &= s(x+1, y) \\
S(x+1, y+1) &= s(x+1, y+1) \\
S(x+1, y-1) &= s(x+1, y-1)
\end{align*}
\]

\[S(x, y) = \frac{1}{4} \sum_{n=0}^{3} s(x+n, y+n) \]

Figure 4. Fractional differential mask in the negative direction of x-axis on numerical computing rules

(d) The limited weighted sum will be performed fractional differential mask respectively as the processing results in 8 directions. That is also the approximation of n-order fractional partial differential of \(s(x, y)\).

(e) To obtain a complete digital image after fractional differential, we should repeat above steps and traverse overall width of the image. So the approximation of n-order fractional partial in 8 directions can be calculated.

During the translation by pixel, in order not to make the line or row of fractional differential out of the image plane, we must keep the distance from image edge pixels to the center point of fractional differential mask less than \(\frac{n-1}{2}\) pixels. That is, \(n-1\) lines or rows of image edge do not need fractional differential.

IV. ANALYSIS ON SIMULATION RESULTS

A. Texture Detail Strengthening

During detailed operations, in order that image fractional order differentiation can have better rotation invariance, this paper simultaneously selects four fractional order differentiation mask operators with \(x, y\), the right diagonal and the left diagonal to respectively calculate image, \((x, y)\) pixel point and its neighborhood area. Then, results of four algorithms are compared and the maximized value is taken as gray value in fractional order differentiation of \((x, y)\). To obtain a clearer improved image, the original image needs to be superimposed between points with its corresponding pixel point gray value of fractional order differentiation. For the fractional order differentiation processing of RGB color image, it needs fractional order differentiate its component R, G, B and synthesize RGB color image. Three fractional mask operators of \(x, y\), the top right corner and the top left corner, selected by following experiments in this paper are shown as figure 5.

From the experiments we know that: first, two-dimensional image fractional order differentiation operator in this paper has strong rotation invariance; secondly, compared to enhancement operator of traditional digital image based on integral order differentiation, fractional order differentiation enhances texture detailed information in smooth region of image, which enhances high-frequency information of image and keeps low-frequency outline information. Towards the image which is rich in detailed texture information, its detailed texture enhancing effect is very prominent.

Since fractional order differentiation of image has powerful detailed texture information enhancement, it can be naturally believed that fractional order differentiation can be applied in biomedical image, medical CT image and satellite remote sensing image. Among them, the experimental results of different enhancement methods on pleural medical image are shown as figures 6.

In order to testify the effect of n-order differentiation mask enhancement on image texture detail, the image obtained by various algorithm processing is respectively vertical projected, to analyze their envelop distribution as figure 7.

From figure 7, we can know that n-order differentiation mask keeps low-frequency outline information to some extent and it reserves the envelop distribution of gray vertical projection on original image, when enhancing high-frequency edge information of original image, which will not cause high frequency of image to be totally white. Compared to traditional image processing algorithms based on integral order differentiation, n-order differentiation mask enhances texture detail characteristics in nonlinearity and it will not cause smooth region to become totally black. So n-order fractional mask has capability of nonlinear enhancement texture detailed information. It can keep integral low-frequency outline information of image and enhance high-frequency edge information with large-range change.
of gray on image. Meanwhile, it also enhances texture detailed information of small-range change degree in image gray.

Figure 5. Comparison of fractional differential and integer differential results.

Figure 6. Comparison of vertical projection of gray-level values of medical images.
TABLE I. QUANTITATIVE COMPARISON OF N-ORDER DIFFERENTIAL ON THREE DIFFERENT IMAGES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Information entropy</th>
<th>Average gradient</th>
<th>Figure</th>
<th>Information entropy</th>
<th>Average gradient</th>
<th>Figure</th>
<th>Information entropy</th>
<th>Average gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>7.6978</td>
<td>7.1551</td>
<td>0.0</td>
<td>7.6386</td>
<td>9.270</td>
<td>0.0</td>
<td>7.6276</td>
<td>9.6719</td>
</tr>
<tr>
<td>0.3</td>
<td>7.6811</td>
<td>7.6328</td>
<td>0.3</td>
<td>7.6636</td>
<td>10.230</td>
<td>0.3</td>
<td>7.5490</td>
<td>9.9204</td>
</tr>
<tr>
<td>0.5</td>
<td>7.6713</td>
<td>9.2093</td>
<td>0.5</td>
<td>7.7155</td>
<td>12.638</td>
<td>0.5</td>
<td>7.6445</td>
<td>12.7391</td>
</tr>
<tr>
<td>0.8</td>
<td>7.7906</td>
<td>16.6780</td>
<td>0.8</td>
<td>7.8134</td>
<td>22.822</td>
<td>0.8</td>
<td>7.7925</td>
<td>21.1791</td>
</tr>
</tbody>
</table>

Figure 7. Comparison of different orders' differential on three different images

B. Comparisons of Different Order Differentiation Mask Operators

In order to testify the effect of image matching, this paper selects a group of target images to be rotated and scaled. Then, it selects fractional mask operators with different sizes to test the matching effect with our improved algorithm under the condition of rotation and scaling. From figure 8 we find that, after fractional order differentiation is processed, the image will become clearer and texture detail will be more obvious. While the processing result with order 0.8 appears strong sharpening effect. To evaluate this effect from quantitative perspective, we adopt information entropy and average gradient as image evaluating parameter. The experimental results are shown in table 1.

As shown in figure 8 and table 1, firstly, the average gradients of all processing methods are improved. It indicates that fractional order differentiation can enhance image definition and improve its quality. Average gradient obtained by fractional order differentiation is twice as much as its original image. Its high average gradient indicates that improved project can obtain better image definition. Secondly, for all processing methods, their information entropies are improved. It indicates that fractional order differentiation can enhance image edge and texture detail. The information entropy obtained by fractional order differentiation is higher than its original image. It indicates that fractional order differentiation can obviously enhance image texture detail and approximate to the best effect under appointed orders. Finally, image (b) and (c) are also analyzed to obtain the same conclusions.

By means of images comparison, we can see that extracting edge information of different order differentiation operators is basically the same. With decrease of the differential order, image edge information becomes clearer and fuzzy parts are obviously improved. It indicates that the noise is reduced with differential order reduction. We can see this method can effectively restrain the noise and it overcomes the disadvantages of traditional edge detection operator on noise sensitivity.

V. CONCLUSION

This paper respectively discusses essential reason and distinct advantage of extracting and enhancing image texture detailed information from two perspectives: signal processing and visual model. It also discusses how to construct image fractional order differentiation and its
numerical realization from two aspects, such as computer digital processing and fractional order differentiation value algorithm. We study the dynamic physical significance of fractional order differentiation and deduce the relationship between fractional order differentiation and classical time-frequency analysis. Then, Gaussian difference receptive field model of image fractional order differentiation is proposed from these aspects in signal processing and neural model of biological vision. Finally, the digital image fractional order differentiation model and its operation rules are proposed. Experimental results of simulation and practical cases show that the enhancing effect of fractional order differentiation on texture detailed information in smoothing region is obviously superior to integral order differentiation operation, for image signal with rich texture detailed information.

REFERENCES


Remote Sensing Classification based on Improved Ant Colony Rules Mining Algorithm

Liu Shuying
The Institute of Information Engineering of XianYang Normal University, Shaanxi, China
Email: lsynate@163.com

Abstract—Data mining can uncover previously undetected relationships among data items using automated data analysis techniques. In data mining, association rule mining is a prevalent and well researched method for discovering useful relations between variables in large databases. This paper investigates the principle of traditional rule mining, which will produce more non-essential candidate sets when it reads data into candidate items. Particularly when it deals with massive data, if the minimum support and minimum confidence are relatively small, combinatorial explosion of frequent item sets will occur and computing power and storage space required are likely to exceed the limits of machine. A new ant colony algorithm based on conventional Ant-Miner algorithm is proposed and is used in rules mining. Measurement formula of effectiveness of the rules is improved and pheromone concentration update strategy is also carried out. The experiment results show that execution time of proposed algorithm is lower than traditional algorithm and has better execution time and accuracy.

Index Terms—Ant Colony; Association Rule; Remote Sensing

I. INTRODUCTION

Data mining, which based on specific algorithms, is a technology for researchers to analyze large amounts of data to find potential knowledge, and widely used in the major fields, makes itself be valuable. Cluster analysis is one of the modes of data mining, which classifies the sample data to different types according to similarity rules, so that in the same type data have high resemblance, different type data become high differential. By clustering, it's easy to describe the characteristic of different types. Rule mining technique is a new frontier technology field, which is a combination of many traditional disciplines and new engineering field and is one of the hot spots in data mining research. Association rule mining technology, in general, refers to find interesting connection (Association) or relationship (Correlation) from a large data set, which finds set of attribute values from the data set, namely the frequent items [1-5]. And then apply these frequent item sets to create rules describing correlation relationship. Based on rule mining of the data, implicit and ambiguous connection rules can be obtained from a database. To obtain these rules can be used as knowledge, or into the knowledge base used to guide future activity, such as marketing, decision support, business management, web design, intrusion detection, etc.

For example, in the field of traditional retail, customer enters the store that seem shopping randomly. In fact its shopping has certain continuity, not only has relation with the shopper's own career, education, life habit, but also there is a certain correlation between the purchased items. The classic example is the beer and diapers. This rule cannot be found by people or it is hard to find, and it often needs to use the computer for processing [6, 7]

In intrusion detection, the extraction of rules and knowledge that exist in the data set, and finding the connection between the data is an important work. Determine normal data and abnormal data, which in conformity with some rules by rule mining. These rules form knowledge rule library to guide the intrusion detection rules match. However, data is huge in the face of intrusion detection. If we take artificially analysis and processing, in order to find the hidden rules or knowledge, it is absolutely feasible. On the one hand, there exists a huge amount of calculation, which human can not achieve. On the other hand, sometimes it is difficult to find deeper hidden rules. Only using the application of computer technology, general rules of related patterns can be found in the invasive data [8-10].

Association rule is a kind of rule, which is described by \( X \rightarrow Y \). That means as long as the data set contains \( X \), it will tend to contain \( Y \). \( X \) is antecedent item set and \( Y \) is consequent item set. The following is an example. 91% of female customers will buy other products of this brand in the purchase of a cosmetic liquid foundation. Among them all, 20% of customers purchase the lipstick, eye shadow, sunscreen and other related products. Here, we call a cosmetics is antecedent item set, the other products of this brand is consequent item set. 91% is confidence of rule and 20% is support of rule. Support is fact scale that supports the rules established and confidence is the degree of trust corresponding to the establishment of rule. Usually against problem of association rules, specify minimum support and minimum confidence as the lowest bottom line. Data set \( D \) is called transaction data set. An item set is \( I = \{i_1, i_2, \ldots, i_n\} \), and subset of \( I \) is transaction. Each transaction has a unique identifier called TID. Data set to rule mining is called transaction data set \( D \), which is made up of a serial of transaction TID.
Swarm intelligence is the interdisciplinary of biology, computer technology and other fields. It means in the nature, due to the interaction and cooperation, the single individual with no intelligent or less intelligent to finalize the complex intelligence mission. In the 1990s, the Italian scholar M.Dorigo, V.Maniezzo, A.Colorni were enlightened by the behavior of ants which could find the shortest path between their home and food source to propose the most important algorithm of swarm intelligence-ant colony algorithm. They use it successfully to solve many combinatorial optimization problems such as traveling salesman TSP, graph coloring problems and so on. As one of swarm intelligent optimization algorithms, the ant colony algorithm has been successfully applied in the remote sensing data classification, provides a broader idea for remote sensing image classification. However, the conventional Ant-Miner algorithm takes long computation time with a slow convergence in mining remote sensing image classification rules. We proposes an new ant colony algorithm based on conversional Ant-Miner algorithm.

Ant colony rule mining algorithm can be better found the optimal solution, is robust, but Ant-Miner algorithm also has some deficiencies caused by the ant colony system defects, such as the streamlining of the classification rules set poor, vulnerable local optimal solution, the calculation time is longer. We proposes an new ant colony algorithm based on conversional Ant-Miner algorithm. Firstly, Improve and perfect the formula is not defined and the definition of unreasonable rules validity measure formula. Then, the conventional Ant-Miner algorithm is modified about the strategy of pheromone update by using new pheromone concentration update item and pheromone evaporation coefficient. Lastly, the mutation operator is introduced in algorithm solve. The new ant colony algorithm accelerates effectively the evolutionary process and shorts the calculation time. This property is called closed down attribute of support. The results indicate that the new algorithm obtains the simpler forms of classification rules and reduce the computation time. Improved Ant-Miner algorithm classification results of the overall accuracy can reach more than 85%, Kappa coefficient above 0.81, the remote sensing image classification, provides a broader idea for remote sensing image classification. However, the conventional Ant-Miner algorithm takes long computation time with a slow convergence in mining remote sensing image classification rules. We proposes an new ant colony algorithm based on conversional Ant-Miner algorithm.

Approximation precise rules is defined as (1).

\[
\text{support}(A) - \text{support}(A \cup B) \leq \epsilon . \quad (1)
\]

A rule generation algorithm is given and formal description is as follows.

Mine_rule()
Input: frequent item set \( L, \epsilon, \text{min}_\text{sup} \).
Output: Approximation precise rule \( ARC \).
\( L_t \leftarrow \text{1-frequent item set}; \)
\( K=1; \)
For each classification attribute \( c \in C \) do
Begin
While \( L_k \neq \text{NULL} \) do
Begin
For each frequent item set \( l \in L_k \) do
If \( \text{sup}(c \cup l) \geq \text{min}_\text{sup} \) then
If \( \text{sup}(l) - \text{sup}(c \cup l) \leq \epsilon \) then
\( ARC \leftarrow (l \Rightarrow c) ; \)
\( C_{k+1} = \text{candidate} - \text{Gen}(L_k) ; \)
Item sets that support is not less than \( \text{min}_\text{sup} \) in the \( L_{k+1} < -C_{k+1} \).
\( k++ ; \)
Endwhile
Endfor.
Endfor.

III. A NOVELSCHEME BASED ON IMPROVED ANT COLONY

In the process of the application of association rule generation algorithm, we need to try every possible rules to determine which is the strong association rules, in order to decrease the amount of tries, ignore the unnecessary try, algorithm can be adjusted according to the following two theorems.
Theorem 1. Item set \( X \), there exists rule \( X \Rightarrow (l - X) \) is not strong rule, for any subset \( X_i \) of \( X \), the generated rule \( X_i \Rightarrow (l - X_i) \) is not strong rule certainly.

Prove: Because \( X_i \subseteq X \), support(\( X_i \)) ≥ support(\( X \)) and the following result sets up.

\[
\text{conf}(X_i \Rightarrow (l - X_i)) = \frac{\text{support}(l) / \text{support}(X_i)}{\text{support}(l) / \text{support}(X)} = \text{conf}(X \Rightarrow (l - X))
\]

\( X \Rightarrow (l - X) \) is not strong rule, then

\[
\text{conf}(X \Rightarrow (l - X)) < \minconf
\]

so

\[
\text{conf}(X_i \Rightarrow (l - X_i)) < \text{conf}(X \Rightarrow (l - X)) < \minconf
\]

sets up and \( X_i \Rightarrow (l - X_i) \) is not strong rule certainly.

Theorem 2. Item set \( X \), there exists rule \( Y \Rightarrow X \) is strong rule, for any subset \( X_i \) of \( X \), the generated rule \( Y \Rightarrow X_i \) is strong rule certainly.

Prove: \( \text{support}(X_i \cup Y) \geq \text{support}(X \cup Y) \)

So the following formula sets up.

\[
\text{conf}(Y \Rightarrow X) = \frac{\text{support}(X \cup Y) / \text{support}(Y)}{\text{support}(X \cup Y) / \text{support}(Y)} = \text{conf}(Y \Rightarrow X_i)
\]

\( \text{conf}(Y \Rightarrow X) \geq \minconf \)

\( \text{conf}(Y \Rightarrow X_i) \geq \text{conf}(Y' \Rightarrow X) \geq \minconf \) sets up.

\( Y \Rightarrow X_i \) is strong rule certainly.

Ant colony algorithm can finds the optimal solution and has good robustness. But Ant-Miner algorithm has some defects, which are as follows.

Ant-Miner algorithm only considers prediction accuracy of available rule set and does not take into account compactness of the rules. Rules mined are often more and complex. One reason is when the concentration of pheromone is updated, some parameters are set too simply, leading to the final result not necessarily the optimal solution. Rule pruning makes the final rule set is not necessarily the most compact.

Ant system itself is easy to fall into local optima. As the ant system, ants carrying pheromone update generally adopts a natural way. Once an ant chooses a path, regardless of how its effectiveness it will be conducted to update global pheromone according to this rule. In the initial search, ant movement is random and the solution obtained is not necessarily the optimal solution, which makes pheromone of some poor paths enhanced. It causes error guidance for subsequent ants and causes quality of rules low.

Search time of Ant-Miner algorithm is often longer.

That is because in the initial stage of search operators, difference of pheromone on the path is small. While updating pheromone concentration, pheromone evaporation coefficient is usually set to be a simple fixed value, although the global pheromone has been updated, the effect on the next ant choosing path is not obvious. The ant movement is still very random. Especially when the sample data attribute nodes are more, it is difficult to find the optimal path in a short time. Only by searching a long time can making better path has high pheromone concentration than the other paths achieve. As the searching continues, the colony makes differences of the amount of information increase by a positive feedback and finally obtains the optimal path. In addition, using pruning strategy to remove each attribute node has a certain blindness, thereby increasing computing time. Based on the Ant-Miner algorithm, pheromone concentration update strategy is improved and mutation operator is introduced in order to dig out the better rule sets in a short time.

When ants construct a complete path, attribute nodes need to be selected repeatedly. In theory, this choice can be completely random, but blind choice will require a lot of computing time. Usually in solving practical problems, you can construct a heuristic function related to this issue to guide the choice of colony to shorten calculation time.

Heuristic function \( \eta_i \) of condition term \( \text{term}_i \) is defined as (2).

\[
\eta_i = \frac{\log_2 K - \inf oT_{\text{term}_i}}{\sum_{i=1}^{K} (\log_2 K - \inf oT_{\text{term}_i})}
\]

\( \eta_i \) represents value of heuristic function of condition term \( \text{term}_i \), the value of which reflects its classification ability. \( K \) represents the number of category, \( a \) represents the number of attributes and \( b \) represents the number of value of attribute \( i \). \( \inf oT_{\text{term}_i} \) is defined in (3).

\[
\inf oT_{\text{term}_i} = \sum_{w=1}^{K} \left\{ \frac{\text{freq}_w^{\text{term}_i}}{\left| T_w \right|} \times \log_2 \left( \frac{\text{freq}_w^{\text{term}_i}}{\left| T_w \right|} \right) \right\}
\]

\( \left| T_w \right| \) represents the number of record \( \text{term}_i \) included by set \( T_w \). \( \text{freq}_w^{\text{term}_i} \) represents the number of category \( w \) in interval set \( T_w \). Before the first ants choosing the path, the initial pheromone concentration of each attribute node is set to be the same value.

\[
\tau_{ij}(t = 0) = \frac{1}{\sum_{i=1}^{b_j}}
\]

\( \tau_{ij} \) represents pheromone concentration of condition term. \( \text{term}_i \), \( a \) represents the number of attributes and \( b_j \) represents the number of value of attribute \( i \). Subsequently ants choose a path based on value of heuristic function of the path and amount of pheromone. Probability formula of one \( \text{term}_i \) selected as the current path is (5).

\[
P_{\text{term}_i}(t) = \frac{\tau_{ij}(t)\eta_i}{\sum_{i=1}^{b_j} \tau_{ij}(t)\eta_i}
\]

Each selected attribute node will be added to the path, when the path contains all the attribute nodes, and then
select a category node for label, thus, a full path that is an initial classification rule has produced. Measurement of the effectiveness of rule is defined in (6).

\[
Q = 1, \quad FP + FN + TN = 0
\]

\[
Q = \frac{TP}{TP + FN}, \quad FP + TN = 0
\]

\[
Q = \frac{TP}{TP + FP}, \quad TN + FN = 0
\]

\[
Q = \frac{TP}{TP + FN}, \quad TN + FP + FN + TN, \quad \text{else}
\]

In the original Ant-Miner, pheromone update formula is (7).

\[
\tau_i(t+1) = (1-\rho) \cdot \tau_i(t) + Q \cdot \tau_j(t)
\]

When subsequent ants choose the path, generally they select attribute node which has high pheromone concentration and do not choose attribute node with a low pheromone concentration, making the colony easy to converge to the same rule, so that some of the potential rules will be discarded. This paper improves pheromone concentration update method from two aspects. Pheromone concentration function is shown in (8).

\[
\tau_i(t+1) = (1-\rho) \cdot \tau_i(t) + \left( \frac{Q}{1+Q} \right) \cdot \tau_j(t)
\]

Increasing of pheromone will be inhibited to a certain degree, which will not lead to converge to local optima due to fast pheromone increasing. Pheromone evaporation coefficient \(\rho\) is improved. Generally speaking, small \(\rho\) will make pheromone of a few paths aggregate, making more ants follow this path, so searching result is not necessarily optimal solution. Big \(\rho\) makes it difficult to distinguish pheromone on the path from each other in a short period of time, so that it reduces the convergence rate and prolongs the computation time. Therefore the size of \(P\) directly affects quality of search results and computation time of ant colony algorithm. In Ant-Miner algorithm, \(\rho\) is usually set to be a fixed value, which makes ants search for a long time and not easy to get the best rules, thus affecting the classification accuracy and efficiency. When updating pheromone (8) is adjusted adaptively and the initial value of \(\rho\) is 0. When the number of cycles is large and the optimal solution has no significant improvement, take the following parameter adjustment strategy.

\[
1 - \rho(t+1) = \begin{cases} 
0.95(1-\rho(t)), & 0.95(1-\rho(t)) \geq \rho_{\text{min}} \\
\rho_{\text{min}}, & \text{else}
\end{cases}
\]

After multiple cycles, the algorithm has no significant improvement, it can enlarge pheromone evaporation of ant path to inspire subsequent ants to find other paths to increase search range. This can make valid information searched by the front ants to be retained, shorten time to speed up the convergence and can make effective searching in the global scope.

After the initial rule pruning, a single point mutation is performed. Randomly select an attribute node and replace it by another random attribute node corresponding to this attribute to obtain a new rule. If the validity of the existing rules is less than the new rules, do mutation operation. Otherwise keep the original rules. This operation is easy to implement and shorter time-consuming, but it can expand the search scope to improve the quality of search results, which can accelerate the convergence to shorten computation time. The mutation process is as follows. Input is rule A and output is rule B. Randomly select an attribute node \(\text{term}_j\) in A. Rule B = Rule \(A_{\text{term}_j=\text{term}_{k}}\), \(j \neq k\);

\[
\text{IF}(A_q < B_q) \quad \text{RETURN Rule B};
\]

ELSE \text{RETURN Rule A}

END

Rule B = Rule \(A_{\text{term}_j=\text{term}_{k}}\) represents that Rule B is generated using \(\text{term}_k\) to replace \(\text{term}_j\) in A. \(A_q\) represents effectiveness of A and \(B_q\) represents effectiveness of B.

IV. EXPERIMENT AND ANALYSIS

In order to demonstrate the validity of our proposed solution, and verify the efficiency of the designed algorithm, we choose 50% and 100% testing data of 10% testing data set of kddcup99 for the related experiment, and compare it with the Apriori algorithm. Experimental environment are HP computer, AMD CPU 2.0GHz and 2G RAM. Software environment are Windows XP SP2 and MATLAB 10.0.

Fig. 2 is comparison result on running time based on different sup on dataset1 and Fig. 5 is comparison result on running time based on different sup on dataset2. Fig. 6 is comparison result on finding right rules based on different sup on 100% testing data. In Fig. 2 and Fig. 4, the abscissa indicates the support and the vertical axis indicates the running time in seconds. In Fig. 3, Fig. 5 and Fig. 6, the abscissa represents trust degree and the vertical axis represents the number of rules. It can be seen from the comparison in this picture execution time of proposed algorithm is lower than Apriori algorithm. Fig. 3 is comparison result on finding right rules based on different sup on dataset1 and Fig. 4 is comparison result on finding right rules based on different conf on dataset1. Under the condition of same support or trust, the ability to find the precise rules of Apriori algorithm is better than our proposed algorithm. But our proposed algorithm is closed to Apriori algorithm when support value or trust value is high. Under the condition of low threshold, its solving ability is less than Apriori algorithm.

Through the above experimental data contrast analysis, it is known that although our proposed algorithm is less than Apriori algorithm on the ability to find accurate rules. From Fig. 2 and Fig. 5, we can see that operation time of our proposed algorithm is significantly shorter than Apriori algorithm, which is because the Apriori algorithm needs to scan the transaction data set for many times.
In order to further demonstrate classification rule mining advantages of ant colony algorithm and reliability of improved algorithm, the study also uses the classical method, decision tree classification rule mining method for the experiment of two years remote sensing image data. Since Landsat TM/ETM+ data has high ground resolution and spatial resolution, and has wide coverage and short cycle, it is more suitable for the study of remote sensing land cover classification. So this article chooses Landsat TM/ETM+ data for land cover classification based on ant colony algorithm. The main parameter of TM/ETM+ is shown in TABLE I. Experimental data of this paper include Landsat-5 image in October, 2006 and in September, 2009. Track number are 123/40 and 123/41. Images are clear and cloudless. Due to the inherent characteristics of the earth and the structure performance of sensor itself, remote sensing image will produce certain geometric distortion. So remote sensing images should carry out more precise geometric correction before the actual application, the purpose of which is to correct image distortion to achieve geometry integration of remote sensing image and standard image. Compactness of rules is estimated by the number of rules and average number of conditions. Rules mined by improved Ant-Miner algorithm, Ant-Miner algorithm and decision tree are compared, which is shown in TABLE II. When adopting improved Ant-Miner and Ant-Miner algorithm, the number of rules and average condition number is smaller than decision tree method. Improved Ant-Miner has the least rules and average condition, meaning that compactness of proposed scheme is the best.

### TABLE I. THE MAIN PARAMETER OF TM/ETM+

<table>
<thead>
<tr>
<th>Band</th>
<th>Band type</th>
<th>Spectral range(μm)</th>
<th>Spatial resolution(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>blue</td>
<td>0.45-0.52</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>green</td>
<td>0.52-0.60</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>red</td>
<td>0.63-0.69</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>near infrared</td>
<td>0.76-0.90</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>middle infrared</td>
<td>1.55-1.75</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>Thermal infrared</td>
<td>10.4-12.5</td>
<td>120</td>
</tr>
<tr>
<td>7</td>
<td>far infrared</td>
<td>2.08-2.35</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>panchromatic band</td>
<td>0.52-0.90</td>
<td>15</td>
</tr>
</tbody>
</table>

### TABLE II. THE NUMBER OF RULES AND AVERAGE CONDITION OF THREE ALGORITHMS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Number of rules in 2006</th>
<th>Average condition number in 2006</th>
<th>Number of rules in 2009</th>
<th>Average condition number in 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed scheme</td>
<td>16</td>
<td>2.4</td>
<td>9</td>
<td>1.9</td>
</tr>
<tr>
<td>Ant-Miner</td>
<td>25</td>
<td>3.1</td>
<td>19</td>
<td>2.7</td>
</tr>
<tr>
<td>Decision tree</td>
<td>53</td>
<td>7.7</td>
<td>46</td>
<td>5.8</td>
</tr>
</tbody>
</table>

Computation time of improved Ant algorithm and Ant-Miner algorithm is shown in TABLE III. Computation time of improved Ant algorithm for remote sensing image data in year 2006 is 35.62 seconds and computation time of Ant-Miner algorithm is 49.56 seconds. Computation time of improved Ant algorithm for remote sensing image data in year 2009 is 20.92 seconds and computation time
of Ant-Miner algorithm is 38.14 seconds. Computation time of improved Ant algorithm is 13.94 seconds shorter than Ant Miner algorithm and computational efficiency is relatively increased by 39% in year 2006. Computation time of improved Ant algorithm is 17.22 seconds shorter than Ant Miner algorithm and computational efficiency is relatively increased by 82%. Experimental results fully show the improved Ant-Miner algorithm can effectively save calculation time, and greatly improve the efficiency of classification. Remote sensing image data classification results of improved Ant-Miner algorithm in year 2009 is shown in Fig. 7, remote sensing image data classification results of Ant-Miner algorithm in year 2009 is shown in Fig. 8 and remote sensing image data classification results of decision tree algorithm in year 2009 is shown in Fig. 9. The green represents forests, the yellow represents construction land and the blue represents water areas.

TABLE III. COMPUTATION TIME OF PROPOSED SCHEME AND ANT-MINER

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Year 2006</th>
<th>Year 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed scheme</td>
<td>35.62</td>
<td>20.92</td>
</tr>
<tr>
<td>Ant-Miner</td>
<td>49.56</td>
<td>38.14</td>
</tr>
</tbody>
</table>

Remote sensing image classification results generally contain error, at the same time in order to more in-depth analysis of the effectiveness of various methods used for remote sensing classification, the accuracy of classification results should be checked. The commonly used evaluation indexes include producers precision, user accuracy, the overall accuracy and Kappa coefficient. Error matrix of the classification result in year 2009 is shown in TABLE IV. TABLE IV A represents error matrix of the classification of improved ant-miner, TABLE IV B represents error matrix of the classification of ant-miner, TABLE IV C represents error matrix of the classification of decision tree, and TABLE IV D represents error matrix of the classification of maximum likelihood. Type 1 represents water area, type 2 represents construction land, type 3 represents forest land, type 4 represents arable land, and type 5 represents bare land. Total accuracy is 87.8594% and Kappa coefficient is 0.8421, when adopting improved ant-miner, total accuracy is 85.4633% and Kappa coefficient is 0.8108, when adopting ant-miner, and total accuracy is 82.5879% and Kappa coefficient is 0.7730, when adopting decision tree. When adopting maximum likelihood algorithm, total accuracy is 79.3930% and Kappa coefficient is 0.7311. It can be seen that classification results of improved ant-miner algorithm is increased by 8.4664% compared with classification results of the maximum likelihood and Kappa coefficient is increased by 0.111. Classification results of ant-miner algorithm is increased by 6.0703% compared with classification results of the maximum likelihood and Kappa coefficient is increased by 0.0797. Classification results of improved ant-miner algorithm is increased by 2.3961% compared with anti-miner and Kappa coefficient is increased by 0.0313. Classification results of improved ant-miner algorithm is increased by 5.2715% compared with classification results of decision tree and Kappa coefficient is increased by 0.0691. Classification results of ant-miner algorithm is increased by 2.8754% compared with classification results of decision tree and Kappa coefficient is increased by 0.0378. Accuracy of classification results of improved Ant-Miner algorithm is higher than Ant-Miner and decision tree algorithm.

TABLE IV. A) ERROR MATRIX OF THE CLASSIFICATION RESULT OF IMPROVED ANT-MINER

<table>
<thead>
<tr>
<th>Type1</th>
<th>Type2</th>
<th>Type3</th>
<th>Type4</th>
<th>Type5</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>189</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>95.94</td>
</tr>
<tr>
<td>286</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>21</td>
<td>90.79</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>334</td>
<td>46</td>
<td>0</td>
<td>85.42</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>34</td>
<td>220</td>
<td>0</td>
<td>80.97</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>5</td>
<td>81</td>
<td>94.19</td>
<td></td>
</tr>
</tbody>
</table>

B) ERROR MATRIX OF THE CLASSIFICATION RESULT OF ANT-MINER

<table>
<thead>
<tr>
<th>Type1</th>
<th>Type2</th>
<th>Type3</th>
<th>Type4</th>
<th>Type5</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>182</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>93.33</td>
</tr>
<tr>
<td>2</td>
<td>271</td>
<td>7</td>
<td>2</td>
<td>27</td>
<td>87.70</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>329</td>
<td>49</td>
<td>0</td>
<td>83.29</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>36</td>
<td>213</td>
<td>0</td>
<td>79.18</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>9</td>
<td>75</td>
<td>89.29</td>
<td></td>
</tr>
</tbody>
</table>

C) ERROR MATRIX OF THE CLASSIFICATION RESULT OF DECISION TREE

<table>
<thead>
<tr>
<th>Type1</th>
<th>Type2</th>
<th>Type3</th>
<th>Type4</th>
<th>Type5</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>179</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>94.21</td>
</tr>
<tr>
<td>2</td>
<td>272</td>
<td>13</td>
<td>8</td>
<td>3</td>
<td>82.93</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>318</td>
<td>58</td>
<td>0</td>
<td>81.33</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>42</td>
<td>198</td>
<td>2</td>
<td>74.16</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>9</td>
<td>67</td>
<td>88.16</td>
<td></td>
</tr>
</tbody>
</table>

D) ERROR MATRIX OF THE CLASSIFICATION RESULT OF MAXIMUM LIKELIHOOD

<table>
<thead>
<tr>
<th>Type1</th>
<th>Type2</th>
<th>Type3</th>
<th>Type4</th>
<th>Type5</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>178</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>92.71</td>
</tr>
<tr>
<td>5</td>
<td>270</td>
<td>13</td>
<td>8</td>
<td>41</td>
<td>80.12</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>305</td>
<td>67</td>
<td>0</td>
<td>79.02</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>52</td>
<td>187</td>
<td>7</td>
<td>69.00</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>54</td>
<td>81.82</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

This paper investigates the principle of association rule mining. Ant colony rule mining algorithm can better find the optimal solution and is robust, but Ant-Miner algorithm also has some deficiencies caused by the ant colony system defects, such as the streamlining of the classification rules set poor, vulnerable local optimal solution and the calculation time is longer. We propose a new ant colony algorithm based on conventional Ant-Miner algorithm. Firstly, Measurement formula of effectiveness of the rules is improved. Pheromone concentration update strategy is carried out in terms of pheromone concentration increasing item and pheromone evaporation coefficient. Lastly, the mutation operator is introduced in algorithm. The new ant colony algorithm accelerates effectively the evolutionary process and shorts the calculation time. The experiment results show that execution time of proposed algorithm is lower than Apriori algorithm and has better execution time and accuracy. It can be used in intrusion detection area and remote sensing image processing area.
ACKNOWLEDGE

This study was supported by Xianyang Normal University Research Fund (Grant No. 13XSYK054).

REFERENCES

[16] L. Dridi, M. Parizeau, A. Mailhot and J. P. Villeneuve, “Using evolutionary optimisation techniques for


Automatic Target Detection Method for High-Resolution Synthetic Aperture Radar Images

Xu Hui and Gu Hong
Nanjing University of Science & Technology; Jiangsu Nanjing 210094, China
Email: 15150646506@163.com

Abstract—Automatic target detection is of great importance in high-resolution synthetic aperture radar (SAR) images processing. In this paper, we proposed a hybrid HMM-TSVM model to detect targets in SAR images. Our proposed SAR image target detection system is made up of three steps. In this first step, the testing/training SAR images are pre-processed, and image visual features are extracted through 2DPCA, which is an improved version of principal component analysis. In the second step, we use the HMM model to construct the training sequence from training image dataset. Particularly, the relationships between the image in the azimuth and feature vectors are used to generate the feature sequences and training sequences for hidden Markov model. Furthermore, the feature sequences and training sequences are extracted from feature vectors with similar target images in an azimuth. In the third step, targets are detected from testing SAR images by TSVM classifier based on the training sequence by exchanging the labels of pair of different unlabeled samples to solve an objective function. Experimental results demonstrate the effectiveness of the proposed algorithm.

Index Terms—Target Detection; SAR Image; HMM; TSVM; 2DPCA

I. INTRODUCTION

Synthetic aperture radar (SAR) refers to a form of radar which is exploited to create images of an object. Synthetic aperture radar denotes an active microwave remote sensing imaging radar, which can obtain data in all weather conditions and has been widely applied in ocean surveillance and monitor [1] [2]. For example, space-borne SAR can be exploited to detect the ocean oil pollution and seaport state timely.

Synthetic Aperture Radar images are usually collected by various platforms (for example the U2, Global Hawk, or F-15E) for different various purposes, such as reconnaissance and targeting [3]. The collection capacity for such imagery is growing rapidly, and along with that growth is the expanding need for computer-aided or automated exploitation of SAR images.

SAR utilizes the motion of the SAR antenna over a target region to give accurate spatial resolution. SAR is typically set on aircrafts or spacecrafts. Particularly, SAR is developed from an advanced mode of side-looking airborne radar [5-7]. Moreover, SAR can create high resolution images with comparatively small physical antennas whether using physical aperture or synthetic aperture.

In order to obtain a SAR image, successive pulses of radio waves should be transmitted to describe a target scene. Next, the echo of each pulse is received and saved as well. Afterwards, the pulses are transmitted and the echos received through a single beam-forming antenna, with wavelengths utilized within the range from a meter to several millimeters [8]. Considering that the SAR devices move with the aircraft or spacecraft, the antenna location which is related to the target will change as well.

As is shown in Fig. 1, three distinct types of ground vehicles (BMP, BTR70, and T72) are given. For each category, optical image and the related SAR image are illustrated.

As is well known that SAR images have been widely used in remote sensing and other applications [9-11]. SAR can also be implemented as inverse SAR by detecting a moving target for a constant substantial time with a specific antenna.

One aspect of the aided/automated exploitation is automatic target detection. Particularly, automatic target detect from synthetic aperture radar (SAR) image is an important aspect of current computer vision research. Automatic target detection algorithm aims to seek targets such as regions in a SAR image and then calculates a class for each region of interests. Automatic target detection in SAR image usually involves training process with example images of known objects [12-15].

In recent years, automatic target detection for SAR images has been a hot research topic in computer vision field. In the following parts, some typical works about target detection for SAR images are listed as follows.

Amoon et al. proposed a novel method for automatic target detection in (SAR) images. In this algorithm, a genetic algorithm-based feature selection and a SVM classifier are utilized to choose the optimal feature subset of Zernike moments for reducing the computational complexity. Particularly, the feature extraction using Zernike moments has the characteristics of linear transformation invariance and robustness [9].

Choi et al. proposes an efficient approach of classifying targets by utilizing a classifier level fusion of range profile and inverse synthetic aperture radar image, at the same time, a scenario-based construction approach of the training database is proposed as well. In particular, the range profile and the inverse synthetic aperture radar image can be exploited in classifying unknown targets [10].
Qin et al. presents an optimization algorithm based on the maximal likelihood estimation to obtain the model for jointly capturing target motion and scattering parameters. Afterwards, initial value calculating approaches with targets’ ghost images can be presented to ensure global and rapid convergence [11].

Pan et al. concentrated on the modulation effect of rotationally symmetric ballistic targets with precession, and then they utilize Doppler in the radar backscattering and revealed by short-time Fourier transform [12].

Wang et al. focused on detect the ship targets in high-resolution inverse synthetic aperture radar image using the complex motion. Furthermore, the authors proposed a novel SAR target algorithm based on the extension form of match Fourier transform, which can compute the parameters of multi-component cubic chirps through 2D maximisations and Fourier transform [13].

In paper [14], sun et al. proposed a novel SAR target detection model to find the rigid body micro-motion target by rotating parts, and then obtains the formulas of micro-Doppler based on target rotation. To promote the accuracy of SAR target detection, low frequency filter algorithm is proposed to separate the echoes of the rigid body [14].

Wang et al. proposed a novel method to detect moving targets in SAR, in which a new forward model based on the temporal Doppler induced by the movement of antennas and moving targets as well.

Particularly, the authors construct the reflectivity images of the moving targets and then calculate the motion parameters based on a filtered back projection technology which can be integrated with the contrast or gradient optimization approach [15].

However, the above SAR targets detection methods have not effectively utilized global features and local features in SAR images. On the other hand, the relationship between the image in the azimuth and feature vectors has not been fully utilized. Therefore, the target detection results in the existing methods can not be satisfied. Compared with the former researches, the main innovations of this paper lie in the following aspects:

1) In the testing/training SAR images pre-processing, we introduce 2DPCA in image visual feature extraction.

2) Hidden Markov model is used to build the training sequence from the visual features in training SAR images, which can represent a SAR image as a feature vector based on training sequence, and then convert the target detection problem to a classification problem.

3) In the classification process of the proposed algorithm, transductive support vector machine is utilized, which can fully use the unlabeled samples.

This paper illustrates a novel automatic target detection method for high-resolution synthetic aperture radar images. The rest of the paper is organized as follows. Section 2 illustrates the SAR image pre-processing and feature extraction. In section 3, HMM based training sequence modeling is given. Section 4 proposes the SAR image target detection method from the HMM based training sequences. To demonstrate the effectiveness of the proposed algorithm, experiments are conducted in section 5. Finally, the conclusions are drawn in section 6.

II. SAR IMAGE PRE-PROCESSING AND FEATURES EXTRACTION

As is shown in Fig. 1, flowchart of the proposed SAR image target detection method is illustrated in Fig.1. The main ideas of our method lie in three aspects: 1) The testing/training SAR images should be pre-processed in advance, and then image visual features are extracted based on 2DPCA, 2) We utilize the HMM model to construct the training sequence from the visual features in training SAR images, and 3) Exploiting the proposed HMM based training sequence model, testing SAR images are classified by TSVM to detect targets. In this section, we explain how to implement the SAR image pre-processing and feature extraction.

![Flowchart of the proposed SAR image target detection method](image_url)
Because the original SAR images in the MSTAR dataset is quite complex, it is necessary to make the preprocessing to implement the image segmentation in advance. Firstly, the original images are transformed by the logarithm conversion. This process could convert speckles from the multiple model to additional model. Secondly, for each pixel \((x, y)\) if a SAR image \(I\), the mean \(\mu\) and the variance \(\sigma\) can be calculated by the following equation.

\[
\begin{cases}
(x, y) \in T, & T(x, y) = 1, \text{if } I(x, y) > \mu + \lambda \sigma \\
(x, y) \in B, & T(x, y) = 0, \text{otherwise}
\end{cases}
\]  

(1)

As the image enhancement technology can prune some information with no use, we exploit the power-law transformation to implement the image enhancement process for the target image as follows.

\[
L(x, y) = \left[ F(x, y) \right]^\beta
\]  

(2)

where \(L\) and \(F\) represents the latter and former transformed image respectively, and \(\beta\) refers to a constant. Furthermore, in the real case, the distance between the SAR image and the original image is quite different, hence, it is of great importance to normalize the target image.

\[
N(x, y) = \frac{L(x, y)}{\sqrt{\sum_i \sum_j \left| F(x, y) \right|^\beta}}
\]  

(3)

where \(N(x, y)\) denotes the normalized image. After the image enhancement, the method of features extraction for SAR image is given as follows.

In this paper, an improved principal component analysis method is used to extract features from SAR images. The method we used is named two dimensional principal component analysis (2DPCA). For a given training set \(\{X_1, X_2, \ldots, X_n\}\), 2DPCA utilizes training images to build the covariance matrix as follows.

\[
C = \sum_{i=1}^{n} (X_i - \overline{X})^T (X_i - \overline{X})
\]  

(4)

where \(X\) refers to the \(i\)th training image with a \(m \times n\) matrix, and \(\overline{X}\) represents the mean sample matrix of all training sample matrix. \(N\) means the number of training images. In order to obtain high quality projection vectors, the following equation should be utilized.

\[
J(w) = \text{tr} \left[ E \left( (X - E(X))w \right)^T (X - E(X))w \right]
\]  

(5)

Based on the covariance matrix defined in Eq.4, the following equation can be obtained as follows.

\[
J(w) = w^T C w
\]  

(6)

Actually, the optimal projection axes \((w_1, w_2, \ldots, w_d)\) refers to the orthonormal eigenvectors of the covariance matrix \(C\). Afterwards, for each training sample, the feature \(F = [f_{i1}, f_{i2}, \ldots, f_{id}]\) can be calculated by the following equation.

\[
f_{ik} = X_i \cdot w_k, k \in \{1, 2, \ldots, d\}
\]  

(7)

Similarly, the feature matrix of each testing SAR image can be calculated by the 2DPCA.
III. HMM BASED TRAINING SEQUENCE MODELING

Hidden Markov model (HMM) is distinguished from a general Markov model in which the states in an HMM could not be estimated directly and can not be observed based on another set of stochastic processes \cite{20,25}. Supposing that total number of states is \( N \) and number of distinct observations for each state is \( M \). Moreover, \( S = \{s_1, s_2, \ldots, s_N\} \) and \( V = \{v_1, v_2, \ldots, v_M\} \) represent the state set and the observation set respectively. Furthermore, an hidden Markov model \( \lambda \) is defined as \( \lambda = (\pi, A, B) \), where \( \pi \) refers to a vector of the probabilities for the initial state, \( A \) means the transition matrix to record the probabilities of transition between states, and \( B \) represents a matrix of the observation probability densities.

 Afterwards, the feature sequence and HMM training sequence generating process is described in Fig.2. We use the relationships between the image in the azimuth and feature vectors to generate the feature sequence and training sequences for hidden Markov model. The feature sequences and training sequences are obtained from feature vectors with similar target images in an azimuth. Each kind of target sample contains a large number of images with different azimuth, and the azimuth angle is in the range of 0 degree to 360 degree with the interval is one degree. Next, we extract the feature vectors in a same category, and then rank the feature vectors according to the azimuth of each image. Hence, the feature vector sequence is obtained, in which each element in a specific feature vector can form a sequence.

 Then, we use the feature vector of a specific image as a starting point, and choose the \( m \) feature vectors from a feature vector sequence in a rectangular window. Hence, a group of feature sequences with the length of \( m \) is obtained. The training sequence group of HMM can be constructed by connecting all the feature sequences together according to azimuth.

 Therefore, each SAR object image can use feature vectors of a series of adjacent azimuth images to generate a group of feature sequences. As feature sequences are generated by the feature vectors of the multiple images belonged to the same category, it contains the information of multiple images within a same class.

 The distances between observation and hidden states are defined as follows.

\[
D_{ij}(k) = \frac{\alpha}{\gamma(i)} \cdot \exp \left( -\frac{d(v_i,s_j)}{\gamma(i)} \right)
\]  

where \( d(v_i,s_j) \) refers to the Chi-square histogram distance between image visual feature \( v_i \) and hidden state \( s_j \). Particularly, parameter \( \gamma(i) \) can be calculated as follows.

\[
\gamma(i) = \sum_{v_i \in v_c} \frac{d(v_i,s_j)}{n_i}
\]

Assuming that the training sequence \( Q \) is made up of \( \ell \) sequences is defined as follows.

\[
F(Q) = \left\{ \frac{\log P(Q | \theta_1)}{L_1}, \frac{\log P(Q | \theta_2)}{L_2}, \ldots, \frac{\log P(Q | \theta_{\ell})}{L_{\ell}} \right\}
\]  

where \( \theta_j \) refers to the HMM model estimated from sequence \( Q_j \), and \( L_{\ell} \) represents the length of the sequence \( Q \).

IV. TARGET DETECTION FROM SAR IMAGES USING THE HMM BASED TRAINING SEQUENCES

Using the proposed HMM based training sequence modeling method, the SAR image can be represented as a feature vector based on training sequence, and then the rest task we should do is to make classification. The classifier we used is the transductive support vector machine \cite{16,19} (TSVM), which is more powerful than the standard SVM. TSVM integrates SVM with the transductive learning algorithm. For a given set of independent distribution, the labeled samples \( L \) are defined as follows.

\[
L = \{(x_i,y_i), \cdots, (x_n,y_n)\}
\]

s.t. \( x_i \in \mathbb{R}^d, y_i \in \{-1,1\}, i \in \{1,2,\ldots,n\} \)  

Meanwhile, the unlabeled samples based on the same distribution are defined as follows.

\[
U = \{x_1, \cdots, x_u\}
\]

Afterwards, the optimization process of TSVM based on the transductive learning is defined in Eq.13.

\[
\min \frac{w^T w}{2} + C \cdot \sum_{i=1}^{\xi_i} + C' \cdot \sum_{i=\xi_i}^{\xi_i^*} |y_i| \quad \text{s.t.} \quad y_i \left( w^T \phi(x_i) + b \right) \geq 1 - \xi_i, \xi_i \geq 0, i \in L
\quad \left| w^T \phi(x_j^*) + b \right| \geq 1 - \xi_i^*, \xi_i^* \geq 0, j \in U
\]

where \( \xi_i \) and \( \xi_i^* \) are the slack variables of the labeled and unlabeled training samples respectively, and \( C, C' \) refer to user-defined parameters for labeled and unlabeled samples respectively. Based on the above definition, the TSVM algorithm we use is given as the following four steps.

Step 1: Set the parameters \( C, C' \), and the condition \( C > C' \) is satisfied, and then implement an initial SVM classifier based on the labeled samples.

Step 2: Set the number of positive labeled samples from the unlabeled samples utilizing the proportion of positive samples in the set of samples which have been labeled.

Step 3: Classifying all the unlabeled samples with the initial classifier, and then label the unlabeled samples to separate positive and negative ones.

Step 4: Exchanging the labels of pair of different unlabeled samples to solve the objective function in
Eq. 13. Particularly, this process is repeated until there is no pair satisfying the given condition.

V. EXPERIMENT

In this section, experiments are conducted to make performance evaluation based on the MSTAR dataset, which is a SAR radar image database to classify eight kinds of tank targets. Each image has size of about 128 by 128. Moreover, different target images may have slightly different sizes. The images are already centered with 0 degree to 360 degree of orientation. As all the targets have similar length and width, the size and shape of targets are nearly similar to each other. Different form optical images, a SAR image can reflect the internal structure of target’s point scatters. Therefore, SAR images which contain the same target with different orientation angles show high degrees of differences.

The training and testing datasets are described in Table II.

<table>
<thead>
<tr>
<th>Training dataset</th>
<th>Number of SAR images</th>
<th>Testing dataset</th>
<th>Number of SAR images</th>
</tr>
</thead>
<tbody>
<tr>
<td>T72sn_132</td>
<td>232</td>
<td>T72sn_132</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T72sn_812</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T72sn_s7</td>
<td>191</td>
</tr>
<tr>
<td>BTR70sn_c71</td>
<td>233</td>
<td>BTR70sn_c71</td>
<td>196</td>
</tr>
<tr>
<td>BMP2sn_9563</td>
<td>233</td>
<td>BMP2sn_9563</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BMP2sn_9566</td>
<td>196</td>
</tr>
</tbody>
</table>

In experiment scheme 1, we test the effect of logarithm and paw-law transformation with different exponents in the SAR image pre-processing step, and the detection rate with different value of parameter $\beta$ (shown in Eq.2) changing $r$ (shown in Eq.2).

![Figure 3](image.png)  
**Figure 3.** Detection rate when the value of parameter $r$ changing.

In Fig. 3, parameter $r$ represents the number of columns and rows in the identity matrix, and we can know that utilizing the 2DPCA, the highest detection rate is achieved when $\beta$ is equal to 4.

Afterwards, utilizing different target SAR images collected from the MSTAR database, we test the performance 2DPCA for feature extraction with other related methods, that is, PCA [26] and ICA [27], and Experimental results are shown in Fig. 4.

From Fig. 4, we can see that 2DPCA performs best in the above three methods. The reasons lie in that 1) The PCA approach require many more coefficients to represent the feature matrix of a SAR image, 2) For the method of ICA, more computation cost are needed, and 3) Although 2DPCA follows a global feature extraction process which does not suitable to save important local features, this method is more efficient than others.

![Figure 4](image.png)  
**Figure 4.** Detection rate comparison when number of features changing.

In experiment scheme 2, to test the influence of image rotation to the proposed algorithm, we constructed another two datasets (containing seven categories of SAR images) to make performance evaluation. One dataset utilizes SAR images without rotation, and the second one applies SAR images with rotation. Details of these two datasets are illustrated in Table III.

<table>
<thead>
<tr>
<th>Training set</th>
<th>Testing set</th>
<th>Without rotation</th>
<th>With rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1: BTR_60</td>
<td>171</td>
<td>85</td>
<td>171</td>
</tr>
<tr>
<td>C2: 2S1</td>
<td>200</td>
<td>99</td>
<td>200</td>
</tr>
<tr>
<td>C3: BRDM_2</td>
<td>199</td>
<td>99</td>
<td>199</td>
</tr>
<tr>
<td>C4: D7</td>
<td>200</td>
<td>99</td>
<td>200</td>
</tr>
<tr>
<td>C5: T62</td>
<td>200</td>
<td>99</td>
<td>200</td>
</tr>
<tr>
<td>C6: ZIL131</td>
<td>200</td>
<td>99</td>
<td>200</td>
</tr>
<tr>
<td>C7: ZSU_23_4</td>
<td>200</td>
<td>99</td>
<td>200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Training set</th>
<th>Testing set</th>
<th>Without rotation</th>
<th>With rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>83</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>96</td>
<td>2</td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>94</td>
<td>2</td>
</tr>
<tr>
<td>C4</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C5</td>
<td>1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>C6</td>
<td>4</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>C7</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Training set</th>
<th>Testing set</th>
<th>Without rotation</th>
<th>With rotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>84</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>88</td>
<td>2</td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>C4</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C5</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>C6</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>C7</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Table V. Confusion matrix for HMM+TSVM classifier using the dataset with rotation.
From the experimental results in Table 4 and Table 5, it can be seen that the average detection accuracy of the proposed algorithm using the dataset without rotation and with rotation are 88.4%, and 90.6% respectively. Hence, the proposed algorithm can effectively detect targets with high precision whether the SAR image is permitted to be rotated or not.

In the following part, we will test target detection accuracy using the proposed HMM+TSVM with HMM+SVM, and the dataset we utilized is listed in Table 3 (shown in Fig. 5).

![Detection rate](image)

**Figure 5.** Detection rate for two different methods using the dataset without rotation

![Detection rate](image)

**Figure 6.** Detection rate for two different methods using the dataset with rotation

Combining Fig. 5 and Fig. 6, the conclusions can be drawn that our proposed HMM+TSVM performs better the HMM+SVM. The reason lie in that TSVM can effectively promote the generalization of SVM by integrating the information of unlabeled samples, and then enhance the classification accuracy.

Based on the above experimental results, we discover that the proposed algorithm can obviously promote the detection rate of SAR image target recognition than other methods.

VI. CONCLUSIONS

This paper proposes a novel SAR image target detection algorithm by integrating HMM model and TSVSVM classifier. Firstly, the testing/training SAR images are pre-processed, and image visual features are extracted through 2DPCA. Secondly, we utilize the HMM model to construct the training sequence from training image dataset. Thirdly, targets are detected from the testing SAR images by TSVSVM classifier based on the training sequence by solving an objective function.

REFERENCES


Emotion Model of Crowd in Mass Incidents

Yin Yanjun
Nanjing University of Science and Technology/School of Computer Science and Technology, Nanjing, P. R. China

Tang Weiqing
Chinese Academy of Sciences/Institute of Computing Technology, Beijing, P. R. China

Abstract—The paper mainly studies the problem of the forming of group emotion. A new emotion model of crowd based on emotional contagion is proposed. Based on the model, the primary factors of group emotion formation, which are the network structures of interpersonal relationship and the level of emotional dependence between individuals, are analyzed. The evolution process of crowd emotion is simulated under four types of complex network. The simulation results show that the network structure of interpersonal relationship affects the convergence time of group emotion, for example, it takes more time to converge individual's emotions in the regular network than in the other three networks. Not only does the level of emotional dependence on individuals have the significant influence on the intensity of group emotion, but on the formation time of group emotion as well. This study shows that the dynamics of group emotion varies due to the level of emotional dependence and the distinct social network structure of interpersonal relationship.

Index Terms—Group Emotion; Emotion Model; Network Structure; Interpersonal Relationship

I. INTRODUCTION

Group emotion, an important research category of social psychology, draws more and more attentions by multidisciplinary along with the development of group behaviors’ research. Group emotion refers to the moods, emotions and dispositional affects of a group of people. It can be seen as either an emotional entity influencing individual members’ emotional states or the sum of the individuals’ emotional states. The critical factor in the forming of group emotion is emotional contagion. In social interaction, emotional contagion is defined as a process that a person will unconsciously or consciously imitate other people's emotional expressions and experience those emotions [1]. In 19th century, Gustavo Le Bon [2] discussed that group emotion was induced by collective mind in his book The Crowd: A Study of the Popular Mind, and explained further that collective mind would emerge in a crowd of people and influence human behaviors which would not predicted by simply studying one person. Sociological researches showed the relative deprivation is the most important root causing the negative emotion of the person in mass events; it believed that the people in crowd would feel discontent when they compare their positions with others and realize that they are lower than others [3]. Researchers of inter-group emotion theory [4] also pointed out group emotion plays the important role in revealing the mechanism of mass incident and inter-group conflict.

Mass incidents occur frequently due to small conflict among the people; however, these incidents usually cause extremely seriously damage to people’s live, property safety and social stability. At present, scholars have been using computer modeling to simulate individuals’ behavior in mass incidents. Joshua M. Epstein [5] devised a simple Cellular Automata model for simulating civil violence. Feng Pingyao and Wang [6] analyzed the crowd behavior of mass incidents without organization and direct interests based on the work of Epstein. Bu Fanliang and Wang Yiyi [7] presented an affective computing model which simulated the mass incidents with venting anger. For these researches of mass incident, the importance of individual’s emotion for the choice of individual’s behavior is noticed, but the dynamic of group emotion is not discussed, especially the emotional contagion happened in mass incidents. If the group emotion is ignored, the model will not fully reflect the characteristics of crowd behaviors of mass incidents.

In mass incidents, people make choices not only by evaluating the consequences and their probability of occurring, but also and even sometimes primarily at emotional level. At high emotional intensity, there are discontinuous “jumps” in the behavior. Emotional contagion is one of the main mechanisms for people experiencing the high emotion. Meanwhile, it is one of the main factors to influence individuals’ emotion. Hence, the research of dynamic of group emotion based on emotional contagion has fundamental significance for us to understand and simulate the collective behavior in mass incidents.

In this paper, a new method of building the emotion model of crowd based on the works of these scholars from VU University is proposed. In our method, new rules of emotional interaction between individuals are built in line with the degree of personal similarity, and also the effect of different network structures of social relationship on the formation of group emotion is analyzed. The main goal of this paper is to explore the effect of the personal similarity under different interpersonal relationships on the evolution trend of emotional contagion, and to reflect deeply the evolution of group emotion in mass incidents.
II. RELATED WORK

Some authors have discussed how to simulate the emotion dynamic of people. Emotional contagion has something in common with the prevalence of infectious disease on kinetic mechanism; hence many scholars used the epidemiological approach to simulation the dynamic of emotional contagion.

Chao W M and Li [8] assumed that there are three states for an individual in the process of emotional contagion: clean, latent and infected. When exposed to a contagion environment where the infected agents in the view exist, a clean agent changes its state to latent. An agent that has been in the latent state for a period of time changes its state to infected. When an individual’s emotion is in the infected state, he has the ability to infect other individuals in emotion. However, the individual has no ability to infect other persons’ emotion if he is in the clean or latent state.

Durupinar [9] provided a threshold model to simulate the spread of emotion, which borrowed the idea from the SIS model. In his model, an individual can be in one of the two states: susceptible or infected. When individuals in the susceptible state come into contact with the infected ones, they can become infected with some probability. Meanwhile, Durupinar introduced the emotional infected dose and emotional threshold in the model to resolve the problem of different ability to infect other persons.

The researchers, from VU University, adopted the head dissipation phenomena studied in thermodynamics to show the process of emotional contagion happened in social interaction, and the Absorption Model [10,11] provided by them is a standard model. In this model, an individual will be infected by his neighbors, and the change of his emotion will depend on his capability to be infected and his neighbors’ abilities to infect him. Mark Hoogendoorn et.al [12] provided an agent-based social diffusion model that simulates the spread of information and emotion among a group of agent based on the method of absorption model. Emotion regulation theory provided by Gross described how a subject can use certain strategies to affect emotion response levels. Based on Gross’ emotion regulation theory, Bosse et al [13] introduced a computational model to indicate how to change certain aspects of behavior or cognitive functioning in order to get a more satisfactory emotion.

Jason Tsai and other researches [14, 15] examined computational models of emotional contagion by implementing VU university and Durupiner, and compared the ability, which provided by two models, to reproduce a real crowd panic. The result showed that the thermodynamics-based model is super to the epidemiological-based model. However, the thermodynamics-based model did not take the social relationships into account, especially the effect of the different network structures of social relationship on the process of emotional contagion of crowd. In fact, research shows that the topological prosperities of interpersonal relationship network considerably affect its function and overall organization [16].

Amount of mass incident literatures showed that the interpersonal relationship between individuals directly affected the emotional intensity of persons in mass incident. From the viewpoint of psychology, the degree of personal similarity is considered to be a form of social distance that influences the decision making and choice behavior of agents [17]. The more similar a pair of agents is, the higher is the propensity (probability) to interact [18] Coeneml R and Broekens J [19] pointed out that interpersonal relationship is one of emotional contagion factors. Goncalo P et al [20] described a computational model designed for emotional contagion simulation in societies of agents, integrating the influence of interpersonal relationships and personality. And there are some scholars introducing the interpersonal factors in the emotional contagion model. However, the interpersonal factor is only represented as a parameter in these models rather than the topological prosperity of interpersonal relationship.

III. MAIN FACTORS OF GROUP EMOTION

A. Interpersonal Relationship Network

Group emotion is the result of the social interaction about cognition, emotion and behavior. In Social interaction, the social relationships between individuals are very important for those persons who transmit their emotions and catch others’ emotions. Social psychology found [21] that the more closely they related to each other, the more obvious the emotional expression is. Hatfield, Cacioppo and Rapson [7] put forward that the process of emotional contagion is unconscious and automatic, and also thought those familiar with others are more likely to transmit their emotional information and more easier to understand others’ emotions to converge them to one emotion.

Interpersonal relationship is usually represented by kinds of graphics that consist of the points (or notes) to represent actors and lines (or edges) to represent ties or relations. To examine the dynamics of the formation of group emotion in different network structures of the interpersonal relationship, we consider four types of complex network: regular network, small world network, random network and scale-free network, to represent different structures of interpersonal relationship.

The regular network is a ring network in which each node is connected with $k$ nearest neighbors. In the paper, each agent is connected to his neighbors ($k=8$) within his Moore neighborhood which is implemented by the square lattice. The method of building the small world network adapted the work proposed by Buncan Watts and Steve Strogatz [22]. It begins with a regular network ($k=8$), and we choose a person and a edge that connects to its neighbors, with rewiring probability $P_r$, and reconnect this edge to other agent who is selected by random, with duplicate edges forbidden. We continue the process until each original edge has been considered once. Then a small world network is built. Increasing rewiring probability $P_r$ until $P_r =1$, all edges are reconnected randomly and the regular network will become a random network. As for the building of Scale-free network [23],
we start with a full-connected network with \( M \) persons, where \( M \) is far smaller than \( N \), the number of the whole of persons. Then additional \( N-M \) persons are added as follow: each new node is linked to \( k (k<=M) \) existing nodes with a probability \( P_e \), that is proportional to the number of links that the existing nodes already have. Formally, the probability \( P_e \) that the new node is connected to node \( i \) is as follow:

\[
P_e = \frac{K_i}{\sum K_j}
\]

where \( K_i \) is the degree of node \( i \) and the sum is made over all pre-existing nodes \( j \).

In the random and WS network, there may be some isolated persons who have no relationship with anyone in interpersonal relationship network. Based on the study of social psychology that emotional contagion can happened between strangers, each agent is connected to his neighbors within his Moore neighborhood to represent the emotional connection between the strangers and to avoid isolating persons. The algorithms of network construction can reference the works provided by Shuguang Suo and Yu Chen [24].

**B. Personality**

Personalities are the psychological characteristics with social significances. Personalities are of intrinsic differences that remain stable throughout most of our life and are the constant aspects of our individuality [25]. One of the most widely accepted models of personality is the FFM in which openness, conscientiousness, extraversion, agreeableness and neuroticism are considered. Based on analyses of emotional contagion psychological mechanism, Verbeke [26] developed a classification method of personality. In line with the ability to infect and the capability to be infected, persons are divided into four different classifications: charismatic, empathetic, expansive and bland.

In the paper, the definition of personality is described as two-dimensional vector based on the work of Verbeke, where each dimension is represented by a personality factor. The distribution of the personality factors is modeled by a normal distribution function \( N \) with mean \( \mu_i \) and standard \( \sigma_i \):

\[
\text{personality} = <\lambda_{Ex}, \lambda_{Es}>
\]

\[
\lambda_i \in N(\mu_i, \sigma_i^2), \mu_i \in [0,1], \text{for } i \in \{\text{Ex}, \text{Es}\}
\]

where \( \text{Ex} \) (Emotional Expressivity) represents the ability to infect other persons through expression, vocalizations and postures, and \( \text{Es} \) (Emotional Susceptibility) denotes the capability to be infected by catching others’ emotion cue, mimicking others’ expression and understanding others’ inner feelings. As an example, an expansive individual that has the high ability to infect and low capability to be infected is represented as personality=<0.8, 0.3>. However, an empathetic individual that has the low ability to infect and high capability to be infected is represented as personality=<0.1, 0.9>, a charismatic individual as personality=<0.7, 0.9>, a bland individual as personality=<0.1, 0.2>.

**C. Intimate Degree**

The influence of interpersonal relationship on emotional contagion is not only reflected by the network structure of interpersonal relationship, but also by interpersonal similarity, trust and other affection factors. The researches of social sharing of emotion found that the lack of trust between individuals will not produce the sharing of emotion. Many realistic examples show that if there is lack of trust or emotional dependence on each other, individuals’ emotion is difficult to change through the unconscious and automatic process of emotional contagion. In the process of social interaction, the intimate degrees between individuals directly influence the individuals’ emotional expressivity and emotional susceptibility. For example, a woman will be very sad when she saw her litter son crying. And if she did not know the child was crying, she would be sad, but the latter level of sad is less than the former one. The research of social psychology pointed out that the expression of emotional behavior is more obviously and strongly when individuals have an intimate relationship. Based on the study of sociologists, a person can easily define the other person who contacts with him as family members, relatives, classmates, colleagues, friends, acquaintances, etc. in line with their intimate degree. However, people’s emotional attachment or emotional trust between them is of difference even they have the intimate relationship. For example, there exists obvious non-equivalence of affectional dependence in father and son. Usually, a child is more emotional dependence on his father, but not vice versa.

According to the above analyses, the intimate degree of interpersonal relationship is defined as \( R \in [0,1] \), which is unequal relationship. In line with the intimate degree of interpersonal relationship, individuals’ interpersonal relationship could be divided into: strong emotional dependence relationship (e.g., family members), and weak dependence emotional relationship (e.g., classmates, friends). If an individual \( i \) has strong emotional dependence relationship with individual \( j \), the intimate degree of relationship is \( R_{ij} \), and \( R_{ji}>R_{ij} \). Normally, \( R_{ij} \) represents the intimate degree which the individual \( j \) emotionally depends on the individual \( i \), and \( R_{ij}\neq R_{ji} \). If individual \( i \) has weak emotional dependence on individual \( j \), the \( R_{ij}>R_{ji} \), where \( R_{ij}\in [0,1] \) is a parameter, constant for every person and period, that controls the whole level of intimacy of individuals.

**IV. EMOTION MODEL OF PEOPLE IN CROWD**

In psychology, emotion is often defined as a complex state of feeling that results in physical and psychological changes that influence thoughts and behaviors. Meanwhile, the behaviors and thoughts of a person or of other persons’ will further influence his emotion, which is called emotion contagion.

Emotional contagion, in conscious and unconscious way to induce or change peoples’ emotion states, is primary mechanism by which individuals’ emotion will
become a group emotion. And the social comparison
theory also gave the interpretations about the creating
group emotion. Individuals tend to compare their
behavior and attitude with others that are like them. In the
process of comparison, individuals coordinate his behaviors, attitudes and emotions with those of others
people.

Based on these analyses and our previous work [27,
28], we develop a simple computational emotion model
visualization the formation process of group emotion. We
consider a population of N interacting agents. For a given
individual A will interact with his acquaintances, that is,
those individuals have the interpersonal relationship with
him. At time step t+1, the emotion intensity of individual
A is E_A(t+1) will change because the emotional contagion
comes from his acquaintances, which is

\[ E_A(t+1) = E_A(t) + \frac{1}{M} \lambda_{E_B}(A) \sum_{B \in R} \Delta E_{AB}(t) \]  

(2)

with:

- \( E_A(t) \) is the emotion intensity of the individual A in the
  last time step t.
- M is the number of individual A’s acquaintances.
- R is the set of his acquaintances.
- \( \lambda_{E_B}(A) \) represents the individual A’s ability to catch
  others’ emotional club and to understand others’ emotional
  behavior.
- \( \Delta E_{AB}(t) \) denotes the emotional intensity change of
  individual A that is caused by the emotional interaction
  with individual B.

If the individual A has the strong emotional
dependence on the individual B and the emotion intensity
of individual B is larger than that of individual A, the
emotion intensity change of individual A is \( \Delta E_{AB}(t) \),
which is

\[ \Delta E_{AB}(t) = \frac{1}{K} (E_B(t) - E_A(t)) \]  

(3)

where K is the average degrees of individuals in the
interpersonal relationship network. The equality of
\( \Delta E_{AB}(t) \) shows the “flow effect” which happened in
the process of emotional contagion, it means that the
individual with high emotion experience will transfer his
emotion to other individuals with low intensity of emotional experience.

If the individual A has the strong emotional
dependence on the individual B, but the emotion intensity
of individual B is less than that of individual A, then the
emotion intensity of individual A is not changed by the
emotional contagion from individual B, which is

\[ \Delta E_{AB}(t) = 0 \]  

(4)

If individual A has the weak emotional dependence
on individual B, the emotion intensity of individual A will
be changed by individual B, which is

\[ \Delta E_{AB}(t) = \lambda_{E_B}(B) (E_B(t) - E_A(t)) \]  

(5)

where \( \lambda_{E_B}(B) \) denotes individual B’s ability to affect the
individual A through emotional behaviors.

V. EXPERIMENTS AND RESULT ANALYSIS

A group emotion model based on emotional contagion
is proposed in the paper. Our aim is to reflect more
believable behavior and emotion of individual in social
situation, and to analyze the influence of the personality
relationship of individuals on the dynamic of group
emotion. A large number of simulations have been
performed to test the model, using simulation software. In
the section, some simulation results are discussed.

A. Network Structure of Interpersonal Relationship

Interpersonal relationship, as the main transmission
medium of emotional information, directly decides the
flow direction of emotional information in the process of
social interaction. Meanwhile, the network structure of
interpersonal relationship is one of main factors which
reflect the characteristic of interpersonal relationship. The
network structure of interpersonal relationship will be
changed with the change of rewiring probability \( P_r \). When
\( P_r = 0.0 \), the network is a regular network, \( P_r = 1.0 \), is a
random network, and while \( 0.0 < P_r < 1.0 \), which has the
characteristic of small-world network.

In this paper, rewiring probability \( P_r \) is from 0.0 to 1.0,
the aim is to explore the effect the network structure of
interpersonal relationship on the group emotion
emergency. The personalties of individual are random
numbers following a normal distribution with mean
\( \mu_{E_B} = 0.7 \), \( \mu_{E_S} = 0.8 \) and standard deviation 0.1, and the
emotion intensities of individuals are generated by a
random number \([0.0, 1.0]\); emotional dependence
threshold \( R_t \) is 0.5. Table I shows the model’s parameters,
the range allowable for each parameter, and the default
values programmed in the simulation unless otherwise
noted.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd size</td>
<td>100 to 1000</td>
<td>100</td>
</tr>
<tr>
<td>Emotional Expressivity</td>
<td>0.0 to 1.0</td>
<td>Normally distributed mean ( \mu = 0.7 ), std.dev 0.1</td>
</tr>
<tr>
<td>Emotional Susceptibility</td>
<td>0.0 to 1.0</td>
<td>Normally distributed mean ( \mu = 0.8 ), std.dev 0.1</td>
</tr>
<tr>
<td>Emotional dependence threshold</td>
<td>0.0 to 1.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Rewiring Probability</td>
<td>0.0 to 1.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Intimate Degree</td>
<td>0.0 to 1.0</td>
<td>Uniform random number in range (0.1)</td>
</tr>
</tbody>
</table>

In this scenario, we consider the rewiring probability
\( P_r \in [0.0, 0.00001, 0.0001, 0.001, 0.01, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0] \)
and the size of population \( N = 289 \). The result obtained is
shown in Fig. 1, 2 and Fig. 3. In these Figures, the X axis
represents rewiring probability \( P_r \), and Y axis represent
the running time of emotion convergence, the average
degree of network or the intensity of group emotion. All
data are the average of 100 tests.

For the group emotion, there is no one common
definition in present, although the concept of a group
emotion has a long history. In this paper, the intensity of
group emotion (GEI) is defined as the average values of
emotion intensity of all individuals in virtual space.
according the “bottom-up” group emotion theory, which is

\[ \text{GEI} = \frac{1}{N} \sum E_i \]  

(6)

where \( E_i \) is the emotion intensity of individual \( i \) in virtual space, \( N \) is the number of the whole of individuals.

As we expected, the network structures of interpersonal relationship, as the media of information flow, play a significant role in the running time of emotional convergence in crowd. From Fig. 1, while rewiring probability \( P_r < 0.001 \) (that is, \( \lg P_r < -3 \)), the average degree of node of interpersonal relationship network is relatively stable, no obvious change. However, while \( P_r > 0.001 (\lg P_r > -3) \), the average degree begins to growing exponentially. And the running time of emotion convergence begins to shorten quickly at \( P_r \approx 0.001 \), see Fig. 2. The result shows that a network with a small average shortest path length and a large clustering coefficient is more conductive for emotional information dissemination when the degree of network is determined. Meanwhile, the larger the average degree of network is, the quickly the emotion convergence is. For example, while \( P_r = 0.1 \) and the average degree of network is 9.51 (which is the average value of 100 runs), the running time of emotion convergence is 285.55, and when the \( P_r \) changed to 0.9 and average degree of network is to 15.80, the running time is changed to 71.03. Because new links can be rewired with the increase of rewiring probability, it makes the average degree of network to be increased and enhances the social interaction of the individuals, which shorts the running time of emotional convergence.

From the experiments, we can draw the conclusion that the network structure of interpersonal relationship has no obvious effect on the intensity of group emotion, see Fig. 3. Although the running time of emotion convergence and the average degree of individuals have changed a lot while rewiring probability \( P_r > 0.001 \) (see Fig. 1, 2), the intensity of group emotion is always about 0.6, see Fig. 3.

**B. Emotional Dependence**

Emotional dependence is built on the interpersonal relationship between persons. Usually, if persons have intimate interpersonal relationship, the emotional feeling of them is strong. For exploring the effect of emotional dependence degree on the formation of group emotion in crowd, we have fixed the emotion levels and personalities of individuals in initial state. The emotional dependence on each other will be controlled by emotional dependence threshold \( R_t \). In the process of emotional interaction, the infected individual A has the emotional dependence on the individual B, If the individual A’s intimacy with individual B is \( R_{AB} \) and is greater than \( R_t \) (\( R_{AB} > R_t \)), the individual A has strong emotional dependence on the individual B. If assume \( R_t \) is equal to \( \beta \in [0.0, 1.0] \) (\( R_t = \beta \)), the ratio of strong emotional dependence relationship to weak emotional dependence relationship is \( 1/\beta - \beta \). If \( R_t = 0.0 \), individuals’ emotional dependence relationships are strong, And \( R_t = 1.0 \), all relationships are weak.

In the experiment, the value of emotional dependence threshold \( R_t \), which decides the level of emotional dependence of all persons in virtual space, is from 0.0 to 1.0 by incrementing 0.1 each time in four types of networks as mentioned above. The size of population is 121 (\( N = 121 \)) The result obtained is shown in Fig. 4, 5 and Fig. 6.

The experiment results show: the larger the emotional dependence threshold, (that is, the weaker the emotional dependence of persons,) the lower the group emotion intensity is under the function of emotional contagion. While \( R_t = 0.0 \), all individuals have strong emotional dependence relationships, and the intensity of group emotion is nearly to 0.97. However, while \( R_t = 1.0 \), all interpersonal relationships of individuals are weak, the intensity of group emotion will decrease to about 0.5, see Fig. 4. Meanwhile, while \( R_t > 0.6 \), it means an person have the intimate (strong) interpersonal relationship with
less than 40%, and others interacted with him are stranger or have weak emotional dependence relationship. In the case, individuals’ emotion will convergent to group emotion in a relatively stable trend; the whole emotion energies of persons did not increase obviously. However, while \( Rt < 0.6 \), the deteriorating trend of individuals’ emotion is very obvious. When \( Rt \in [0.0, 0.2] \), individual’s emotion intensity is nearly up to 1.0 which is the self-limiting of individual emotion. In mass incidents, it is easy for people to create deindividuated when they experience the high negative emotion (the intensity of individuals’ emotion is larger than 0.8). Deindividuation is a concept in social psychology that is generally thought of as the losing of self-awareness and individualism [29] in groups. In mass incidents, deindividuation can increase aggression and other anti-normative collective behavior. Hence, if the members of group are friends or relatives, they will experience the high negative emotion due to emotional contagion in minor skirmish. At high emotional intensity, it is easy to make the minor skirmish evolved into the incident of mass violence.

If the people in mass incidents with strong emotional relationship (\( Rt < 0.2 \)), these people will experience strong negative emotion whatever the network structure of interpersonal relationship is, and it is risky to happen collective behavior. Likewise, when more than 60 percent people have weak emotional relationship (\( Rt > 0.6 \)), the emotion state of people will be relatively normal in large group and the network structure of interpersonal relationship could not exert its function in the process of emotional contagion.

If \( R \in [0.2, 0.6] \), network structures of interpersonal relationship will do his work on the forming of group emotion. The intensity of group emotion is larger in the regular network than other networks, and is smaller in the random network. For example, while \( Rt = 0.3 \), the intensity of group emotion is about 0.65 in the random network, however it is about 0.87 in the regular network, see Fig. 4. The average degree of node in four networks is 12.02(scale-free network), 15.70(random network), 8.0(regular network) and 13.89 (WS network) while rewiring probability \( P_r = 0.5 \). Hence, we can draw the conclusion that the average degree of network in the interpersonal relationship is not the main factor affecting the formation of group emotion.

We found that the higher the percentage of strong emotional dependence is, the larger the intensity of group emotion is, see Fig. 4 and Fig. 5. In Fig. 5, the level of strong relationship (\( L_{SR} \)) is defined as follow:

\[
L_{SR} = \sum N_{SRN} (i) / \sum N_{r} (i)
\]

with:

- \( N_{SRN}(i) \) the number of persons having the strong emotional dependence relationship with the individual \( i \) and locating in Moore neighbor of individual \( i \).
- \( N_r(i) \) the number of persons having the relationship with individual \( i \), that is, it is the degree of node (individual \( i \)) in the interpersonal relationship network.

The location of persons who have the strong emotional dependence relationship decides the intensity of group emotion. For example, when the emotional dependence threshold is 0.3, the level of strong relationship in four networks is about 0.70(regular network), 0.46(scale-free network), 0.40(WS network) and 0.36(random network). The level of strong relationship is largest in the regular network; the intensity of group emotion is up 0.87. Similarly, the intensity of group emotion in other network is according to his level of strong relationship. The result suggests the intensity of group emotion is determined in terms of the level of strong relationship. And it shows the individuals are easy to be triggered their experience high emotion when those people have strong emotional dependence relationship with her / his neighbors.

As for the running time of emotion convergence, it increases at the beginning with the decrease of emotional dependence. However, when it is increase its top, the running time of emotion convergence begins to decrease with the decrease of emotional dependence, see Fig. 5. For example, where \( Rt < 0.2 \) and \( P_r = 1.0 \) (that is a random network), the running time of emotional convergence is about 798 steps, and when \( Rt > 0.2 \), it begins to quickly decrease, and the running time of emotional convergence is down to about 47 steps. The result shows that while strong emotional relationship account for less than 40 per cent of all interpersonal relationship, the individuals’ emotion will quickly converge to one. The running time of emotional convergence will be long while \( Rt < 0.6 \), and it reaches to maximum while \( Rt \) is about 0.2. The result shows that if a group is consisted of individuals who have strong (weak emotional dependence relationship), the running time of emotional convergence will be short relatively.

VI. Conclusions

In the paper, we proposed a computational model of group emotion according to the emotional contagion theory and the works of Bosse et al. Based on this computational model, a number of simulation experiments have been performed.

The model offers several results, some of which are intuitive whereas others are somewhat unexpected. The results show the network structure of interpersonal relationship has a significant effect on the running time of emotional convergence. When the rewiring probability of building interpersonal relationship network is larger than 0.001, the running time of emotional convergence begins to decrease from more than 500 steps down to less than 100 steps. The intensity of group emotion is decided by the level of emotional dependence relationship. If the degree of personal similarity is not considered in the interpersonal relationship networks, these network structures have no obvious affect on the intensity of group emotion. Meanwhile, the locations of persons who have the strong emotional dependence relationship determine the intensity of group emotion while the emotional dependence threshold ranges from 0.2 to 0.6 (\( Rt \in [0.2, 0.6] \)).

The group emotion model provided by this paper embodies the influencing factors of the group emotion formation. The influencing factors contain the degree of...
individual’s similarity and the network structures of interpersonal relationship, which consequently validates the seasonality and effectiveness. For the crowd simulation, the model has some referential values. For example, the group emotion model can be applied to simulating the crowd behavior in virtual meeting and virtual cartoon game.

![Figure 4. Group Emotion Intensity at emotional dependence threshold](image_url)

![Figure 5. Percentage of strong emotional dependence relationship at emotional dependence threshold](image_url)

![Figure 6. Running time of emotion convergence at emotional dependence threshold](image_url)

In the future, we are interested to investigate the relationship between group emotion and choice behavior of individuals in mass incidents. Using the researches of social emotion and individual behavior provided by the Social psychologists, we hope in future we could simulate the crowd behavior under the emotional contagion, specially the violence mass behavior and the deindividuation; The dynamic of individuals’ personality in high negative emotion is one of our main work. And the emotion contagion in massive on line social network is also our future work.

**ACKNOWLEDGMENT**

The authors gratefully acknowledge the support of the scientific research project of institution of higher learning of Inner Mongolia (No. NIZY13047), of the nature science foundation of Inner Mongolia (No. 2014MS0614) and of the scientific research foundation of Inner Mongolia Normal University (No. 2012ZRYB009).

**REFERENCES**


dispersions’. Autonomous Agents and Multi-Agent Systems, 2013, pp. 1-18


[27] Yin Yanjun received the Bachelor's degree from the Inner Mongolia Normal University, and the M.S. degree in computer science from the Inner Mongolia University, Hohhot, China. She is currently an associate professor at the Department of Computer Science and Information Technology, Inner Mongolia Normal University; she is currently working toward the PhD degree in School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, China. She research interests mainly include affective computing and virtual reality.

[28] Tang Weiqing received his PhD degree from the Institute of Computing Technology, Chinese academy of science, Beijing, China, and the M.S. and Ph.D. degrees in computer science from the Nanjing University of Science and Technology, Nanjing, china. He is a member of China Computer Federation, assistant secretary. He has worked at the Institute of Computing Technology, Chinese Academy of Sciences. His research interests include manufacturing information, CAD&CG, virtual reality.
Color Remote-sensing Image Segmentation Based on Improved Region Filter

Hou Lei
School of Information and Electrical Engineering, CUMT. Jiangsu Xuzhou 221000, China
Department of Information and Electrical Engineering, XZIT. Jiangsu Xuzhou 221000 China

Abstract—High resolution remote-sensing images provide abundant color, shape structure and texture information. However, region-based segmentations do not allow to fully exploit the richness of this kind of images. Despite the enormous progress in the analysis of remote sensing imagery over the past three decades, there is a lack of guidance on how to select an image segmentation algorithm suitable for the image type and size. In accordance with the characteristics of color high-resolution remote sensing images, we put forward an improved region filter remote-sensing color-image segmentation algorithm based on traditional JSEG segmentation. In order to well describe the color homogeneity in region, the local homogeneity matrix can be adopted to correct the local value in traditional JSEG segmentation algorithm so as to accurately reflect the boundary region and enhance the accuracy of image segmentation. Simulation experiment results demonstrate that our proposed algorithm can efficiently overcome the inaccurate segmentation of boundary region. In addition, compared with the traditional filter algorithm, our proposed algorithm can extract weak edge more accurately so as to achieve the better segmentation result in high resolution remote-sensing image segmentation.

Index Terms—Remote-Sensing Image; Image Noise; Local Homogeneity Matrix; Over-Segmentation

I. INTRODUCTION

With the rapid development of modern satellite remote-sensing technology, the breadth and depth of the remote-sensing technology has constantly strengthened. Remote-sensing images have different applications in geological science, agriculture, forestry, urban planning, environmental monitoring, military, commercial and other fields. Remote-sensing technology has become one of the key technologies in the realization of digital earth strategic thought [1]. The purpose of remote-sensing image processing is to effectively adopt the remote-sensing data so as to analyze, cluster and interpret the data [2] [3]. Thus, the image data can be converted into useful information so as to solve practical problems [4] [5]. Therefore, fully exploring various spatial information of the image to realize the automatic extraction of the information and automatic identification of the Region Of Interest is one of the main development directions for the remote-sensing digital image processing [6] [7]. Due to the various ground terrain, all kinds of information are covered and interlaced. In addition, since these images still exist some error adopted by different atmosphere and sensor performance, these errors greatly influence the interpretation and extraction of the required information. Only further processing, analysis and interpretation of remote-sensing image data, it can decompose and separate effectively the mixed information. The useless information and noise can be suppressed and ruled out, which shows useful information can be enhanced and highlighted. Therefore, remote-sensing data obtained by remote-sensing technology can be converted into useful information.

Image segmentation [8] is a key step from image processing to image analysis. In addition, it is also a basic computer vision technology [9]. Generally speaking, image segmentation algorithms can be divided into three categories, characteristics threshold or clustering, edge detection and region extraction. In addition to that, there are also hybrid segmentation algorithms [10] [11], such as region growing and region merging, region segmentation and edge detection, the histogram-based and adjacent relation-based fuzzy set theory and neural network algorithm, physical algorithm and hybrid algorithm, etc. Most of the existing image segmentation algorithms are primarily based on gray image, and these algorithms are relatively mature. Most of the gray image segmentation technology can be extended to color image, such as histogram threshold, clustering, region growing, and edge detection algorithm, fuzzy and neural network. In addition to the traditional segmentation algorithm for color images, color images can also be regarded as a multispectral image, so some of the multispectral image segmentation algorithm can be applied to the color image segmentation. Because the remote-sensing images are usually characterized by low contrast, regional characteristics have not fully reliable model and guidance due to different shooting conditions and the complexity and diversity characteristics of shape and texture distribution, which frustrates the development of remote-sensing image segmentation technology. Therefore, there are lots of algorithms for segmenting color remote-sensing images, but no universal algorithm exists. There is not an algorithm to segment the entire remote-sensing images effectively, it is difficult to form a unified standard and commonly adopted criteria.

In view of the above problems, how to well remove noise and reasonably adopt color information to accurately segment image is a hot research field in high resolution remote-sensing image segmentation. In order
to effectively achieve accurate segmentation for region boundary, an improved region filter remote-sensing color-image segmentation algorithm based on traditional JSEG segmentation is proposed in this paper. In order to well describe the color homogeneity in region, the local homogeneity matrix can be adopted to correct the local value in traditional JSEG segmentation algorithm so as to accurately reflect the boundary region and enhance the accuracy of image segmentation. Simulation experiment results demonstrate that our proposed algorithm can efficiently overcome the inaccurate segmentation of boundary region. In addition, compared with the traditional filter algorithm, our proposed algorithm can extract weak edge more accurately so as to achieve the better segmentation result in high resolution remote-sensing image segmentation.

Segmentation principle of color remote-sensing image

II. TRADITIONAL JSEG SEGMENTATION

Remote-sensing image segmentation is realized by using optimal threshold. Image segmentation is just to segment an image into different sub-images so as to get some Regions of Interest (ROI). Main segmentation algorithms include multi-scale algorithm, multi-resolution algorithm, image local-function-based algorithm, boundary curve fitting-based algorithm, Markov random field algorithm, hybrid distribution algorithm, threshold algorithm and watershed transform algorithm, etc. JSEG algorithm [12] [13] is a region segmentation algorithm based on color and texture. In addition, JSEG algorithm can be divided into two steps: color quantization and space region segmentation. Its principle is shown as follows [14] [15].

![JSEG segmentation schematic](image)

Figure 1. JSEG segmentation schematic

Segmentation of color image is a difficult problem in the image processing field. In order to extract the color feature of image and realize the color quantization, the image format is converted from RGB space to LUV space. The image is grayed and then smoothed. Splitting algorithm is first adopted to find out the clustering number. Finally, GLA algorithm is adopted to realize quantization so as to produce “clustering image”. The “clustering image” can be viewed as a special texture image. Space segmentation is to scan "clustering image" with template matching [16] [17]. According to the distribution of color vector in the template, the template center pixel value \( J \) can be calculated so that the regional distribution information of image is reflected by the \( J \) image. In the \( J \) image, the little value represents the center of the region, large value denotes the boundary of the region [18] [19]. According to the initial threshold, a fixed seed region can be selected as the improved threshold in the \( J \) image. Eventually the center of the object area can be obtained [20]. And then, these pixels outside seed region are assigned to the seed region by region growing in accordance with the distance. Finally, the color histogram of the generated object is adopted to perform global optimal region merging so as to complete the final image segmentation [21] [22].

In JSEG algorithm, we first need to know the color histogram distribution of the whole color image so as to perform primitive image segmentation. As we all know, the key technology of clustering image is color quantization. Color quantization technique has been studied for many years. Nowadays, color quantitative algorithms can be divided into 2 categories: segmentation algorithm and clustering algorithm. The basic idea of segmentation algorithm is to regard \( K \) colors with most frequency as a palette, and then the rest of the colors are mapped to palette according to the minimum distance rule. The reconstructed image will have a rich layering, but lose small frequency color. In addition, the detail cannot be retained, which causes local blur. Clustering algorithm is to first choose some clustering-centers, and these color vectors are iteratively clustered by certain rules until the appropriate clustering result is found. Clustering algorithm is an approximate optimal algorithm, but need to be iterative operation with huge computational cost [23] [24]. In addition, quantitative results often rely on the selection of the initial clustering center, and clustering algorithm is easy to merge similar color, which destroys the color layering. At the same time, these algorithms only limit on the analysis of color space, and do not consider the spatial distribution of color. Therefore, it will affect quantitative results [25] [26].

According to human visual characteristics, the color spatial distribution and the minimum quantization distortion, a color quantization algorithm based on \( LUV \) color space is proposed. First of all, parity-group filter is adopted to smooth and remove the noise in the color image, and local statistical properties obtained in the process of filter, are regarded as vector quantization weight. Then, the vector quantization (VP) technology is adopted to cluster the vector quantization result. Research shows that the human visual perception is more sensitive to the change of smooth area than that of the rich details region. Therefore, color quantization can be performed in the detail area without affecting the quality of visual perception. If smooth area is given larger weight and rich details region is given smaller weight, rich details region can have rough quantitative and smooth area can have fine quantitative so as to improve the quality of the whole image. In general, such as color image under the natural scene, the suitable quantitative levels should be set to 10 ~ 20 colors.
Firstly, a \( w \times w \) filter window is selected as operation region. In the window, the Euclidean distance between neighborhood pixel and the center pixel is calculated. Assuming the eigenvector is denoted as \( x_s(n) \), the feature distance between neighborhood pixel and the center pixel in window is written as:

\[
d_i(n) = \| x_s(n) - x_c(n) \|, \quad i = 0, \ldots, k \quad k = w^2 - 1
\]  

(1)

Then, \( d_i(n) \) is performed in an ascending sort order:

\[
d_i(n) \leq d_{i+1}(n) \leq \ldots \leq d_n(n)
\]  

(2)

A. Eliminating Impulse Noise

In the filter window, the first difference \( f_i(n) \) of the Euclidean distance \( d_i(n) \) between neighborhood pixel and the center pixel is calculated. \( f_i(n) \) can be adopted to identify impulse noise and remove it. The first difference expression is shown as follows:

\[
f_i(n) = d_{i+1}(n) - d_i(n)
\]  

(3)

Then, the first \( M \) distance difference values are obtained, \( M = (w+1)/2 \), and the threshold is set to \( \alpha \). If \( \alpha \) satisfies the following inequality:

\[
f_i(n) > \alpha
\]  

(4)

It suggests that the noise exists. Therefore, it will be expatiated as follows:

1. When the first \( M \) points are tested, if they satisfy Equation (4), these points with \( j \leq i \) will be viewed as noise.
2. When the last \( M \) points are tested, if they satisfy Equation (4), these points with \( j > i \) will be viewed as noise.

B. Determining Parity-Group Member

After the above algorithm is adopted to remove noise, parity-group member can be determined according to Fisher criterion. Fisher criterion function is defined as follows:

\[
J(i) = \frac{|a_1(i) - a_2(i)|}{s_1^2(i) + s_2^2(i)} \quad i = 1, \ldots, k
\]  

(5)

where \( a_1(i) \) denotes the average distance of the first \( M \) pixels in filter window, \( a_2(i) \) denotes the average distance of the rest of pixels in filter window. They can be computed as follows:

\[
a_1(i) = \frac{1}{i} \sum_{j=0}^{i-1} d_j(n) \\
a_2(i) = \frac{i}{k+1-i} \sum_{j=i}^{k} d_j(n)
\]

In addition, intra-class divergence degree can be defined as follows:

\[
s_1^2(i) = \sum_{j=0}^{i} \left[ d_j(n) - a_1(i) \right]^2
\]

(6)

\[
s_2^2(i) = \sum_{j=i}^{k} \left[ d_j(n) - a_2(i) \right]^2
\]

(7)

By traversing all the possibilities, when \( J(i) \) is maximum, \( i \) is equal to the size of parity-group. It can be denoted as follows:

\[
m(n) = \arg \max J(i)
\]

(8)

The parity-group member can be written as a collection \( p(n) \), where \( p(n) \) can be defined as follows:

\[
p(n) = \{ x_i(n), i = 0, \ldots, m(n) - 1 \}
\]

(9)

After PGF filter is finished, the parity-group of each pixel in original image has been determined. Therefore, the parity-group information can be adopted for color quantization. Color quantization uses the local statistical characteristics obtained by the parity filter as the weight in the process of vector quantization. Weight mainly relies on local statistical property brought by the parity-group. The main process is shown as follows:

1. According to the number of parity-group member obtained by Fisher criterion, the maximum value between the parity-group and center pixel is selected. Since the Euclidean distance between the neighborhood pixel and the center pixel is in ascending order, so we can define the following equation:

\[
T(n) = d_{m(n)-1}(n)
\]

(10)

It is worth noting that \( T(n) \) denotes the image smoothness in window area. The smaller \( T(n) \) is, the smoother the window image is, whereas window image is rougher.

2. The weight of each pixel can be defined as follows:

\[
v(n) = \exp(-T(n))
\]

(11)

Obviously, since \( T(n) \) is the smaller in the smooth area, so \( v(n) \) is the larger. Otherwise, in complex area, the larger \( T(n) \) is, the smaller \( v(n) \). Therefore, \( v(n) \) denotes the weight of smoothness in the region, and the weight of the noise pixel is less than that of smooth area pixel.

3. Average value \( T_{av} \) of \( T(n) \)

\( T_{av} \) shows the overall smoothness of the image. The greater \( T_{av} \) is, the less smooth the image, which may produce more cluster. \( N \) can be adopted to estimate the number of initial image quantization:

\[
N = \beta T_{av}(n)
\]

(12)

where \( \beta \) is set to 2 in experiment.
Since the goal of GLA algorithm is to minimize the global distortion, there will be multiple clusters in large number of similar color pixels. So the clustering algorithm needs to be adopted to merge similar vector code so as to get the quantitative results. Then, image pixels are marked in color cluster so as to form a new structure, which is called a cluster diagram. Cluster diagram can be considered as a data points set in two-dimensional space; the value of each point is the location of the image pixels in the two-dimensional plane, which is denoted as vector (x, y). These data points have been classified and each point has its own cluster mark. Cluster diagram can be also regarded as a special color texture structure, which denotes the basic color classification within color images.

JSEG algorithm proposes a novel image segmentation criterion of color texture region, which is called as J value. The value is computed on the basis of cluster diagram. It is assumed that Z is a collection of N pixels. Based on the definition of cluster diagram, we know that it is a two-dimensional data structure, where each data point is a one-dimensional vector denoted as (x, y). Let Z = (x, y), z ∈ Z, the average m can be computed as follows:

\[ m = \frac{1}{N} \sum_{z \in Z} z \]  

If Z is divided into C clusters \( i = 1, 2, \ldots, C \), where \( m_i \) is the average of \( N_i \) data points in \( Z_i \) cluster, so we have:

\[ m_i = \frac{1}{N_i} \sum_{z \in Z_i} z \]  

If the following Equations are satisfied:

\[ S_r = \sum_{z \in Z} \| z - m \|^2 \]  

\[ S_w = \sum_{i=1}^{C} S_i = \sum_{i=1}^{C} \sum_{z \in Z_i} \| z - m_i \|^2 \]  

So J can be defined as follows:

\[ J = S_w / S_w = (S_r - S_w) / S_w \]  

This algorithm mainly includes image color quantization and space segmentation. But when the boundary is unclear or different shadows are showed in the same region, the effect of segmentation worsens and the position of boundary will not be accurate. In addition, an over-segmentation phenomenon will appear. Therefore, our paper will focus on the improvement of the traditional segmentation problem in high resolution color remote-sensing images.

III. COLOR REMOTE-SENSING IMAGE SEGMENTATION BASED ON IMPROVED REGION FILTER

It can be seen from the previous sections that JSEG algorithm is an effective color-image segmentation algorithm. However, when the traditional JSEG algorithm is adopted to perform the high resolution remote-sensing image segmentation, the regional boundary location will not be accurate and the same region will occur over-segmentation. Based on the traditional JSEG image segmentation algorithm, a color remote-sensing image segmentation based on improved region filter is proposed. In addition, our proposed algorithm can reflect regional boundary in a better way and improve the accuracy of the boundary segmentation when the local homogeneity matrix is adopted to correct the local value.

In image segmentation, if the extreme value point location is not accurate, it will lead to inaccurate segmentation boundary. In order to solve this problem, the local homogeneity gradient matrix is adopted to correct J value so as to make its extreme value point location more accurate. The value of each pixel is firstly computed to obtain local homogeneity gradient matrix.

Since the pixel points of an image can be seen as a set of data points in a two-dimensional plane, the location of each pixel can be described by coordinates (x, y).

It is assumed that P is a model for computing homogeneity, and the eigenvalue of each pixel (such as intensity, or R,G,B value) is I(x, y). And in general, P is a window with the size \((2N+1) \times (2N+1)\).

\( c = (x_c, y_c) \) is a center pixel of window P, whose feature value is I(x_c, y_c).

The vector \( c_{P} \) is shown as follows:

\[ c_{P} = (x_c - x, y_c - y) \]  

Based on the vector, we can construct a new vector as follows:

\[ \tilde{f}_i = (I(x_c, y_c) - I(x, y)) \frac{c_{P}}{\|c_{P}\|} \]  

Then, in the filter window P, \( \tilde{f}_i \) of all pixels are summed to get \( \tilde{f} \), namely,

\[ \tilde{f} = \sum_{i=1}^{(2N+1)^2} \tilde{f}_i \]  

According to Equation (20), H value of the center pixel c in window P can be defined as:

\[ H = \|\tilde{f}\| \]  

Obviously, the algorithm considers the different between two pixels in more subtle ways. Since the extreme value point of H also reflect the boundary of different region, so we can adopt the H value of each pixel to correct J value so that its extreme position can denote the boundary with more accurate. Our proposed algorithm will be described for further discussion. It is worth noting that the eigenvalue of each pixel can have multiple values for color images, so H values of different eigenvalue will be fused. In this paper, R,G,B
of each pixel in the image are selected as the eigenvalue, and then the \( H \) value of each pixel are defined as follows:

\[
H = \sqrt{H_R^2 + H_G^2 + H_B^2} \tag{22}
\]

Local homogeneity image, namely, \( H \) image, it actually is a gray-scale, where each pixel value is denoted as the \( H \) value. We first define a threshold \( T_{H} \). If the \( H \) value is greater than the threshold, which is marked as white for the border, or marked as black for homogenous region. The below figure shows the high resolution remote-sensing images in Washington and Beijing.

Consider the \( H \) value and local \( J \) value also can adopt extreme value to describe the boundary information of the region. Therefore, this paper proposes the following algorithm to correct local \( J \) values. First of all, Equation (22) can be adopted to get the local homogeneous matrix \( H \) of the high resolution remote-sensing image; and then find out the maximum \( H_{\text{max}} \) in matrix \( H \). \( H \) is normalized, which is shown as follows:

\[
\bar{H}_i = \frac{H_i}{H_{\text{max}}} \tag{23}
\]

Finally, when multiplying \( \bar{H}_i \) by local \( J \) of the pixel, we can get local \( J \) of the new pixel, which is to say that as for each pixel \( p_i = (x_i, y_i) \), the new local \( J \) can be denoted as follows:

\[
J_{\text{new}} = \bar{H}_i \times J_{\text{old}} \tag{24}
\]

In conclusion, the steps of \textit{JSEG} algorithm based on local homogeneity matrix can be expressed as follows:

1) Equation (21) is adopted to get the local homogeneity matrix;
2) Equation (24) is adopted to correct the local \( J \) for each pixel, and get \( J \) image;
3) Region growing is adopted to get final result.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In order to evaluate the segmentation performance of color remote-sensing image based on improved region filter, we make several simulation experiments. All of the experiments are run under MATLAB v7.8 (R2012a) on PCs with an Intel Core i5 CPU at 3.2GHz and 2 GB memory. In this paper, opencv is also adopted to realize the two algorithms. In addition, the image cluster diagram is also merged. The high resolution remote-sensing image of the original Washington is divided into 9 clusters, and then the number of clusters is reduced to 8. The clusters number of high resolution remote-sensing images in Beijing is reduced from 9 to 7. The \textit{JSEG} algorithm and \textit{JSEG} algorithm based on local homogeneity matrix are adopted to segment the two images, respectively. The following experiment results show comparison of the two algorithms under different scales:

![Comparison of J image in big scale](image1)

![Segmentation comparisons in big scale](image2)

<table>
<thead>
<tr>
<th>Table I. COMPARISON OF ALGORITHMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm</td>
</tr>
<tr>
<td>JSEG algorithm</td>
</tr>
<tr>
<td>Improved algorithm</td>
</tr>
</tbody>
</table>

Fig. 2 shows the segmentation precision between our proposed algorithm and the comparison algorithm. From this perspective, our proposed algorithm has relatively good segmentation precision and its segmentation precision is between 89.6%~94.2%; while the precision
of JSEG algorithm is between 76.5%~83.6%. Comparison table of algorithms is shown as follows Table I.

We can see from experimental results that, compared with JSEG algorithm, our proposed algorithm has remarkable effects and high accuracy, and can extract the boundary information, which demonstrates the effectiveness of our proposed method in preserving and recovering image structure, such as edges and textures. In additional, the algorithm can also smooth the effect of noise. The segmentation information is accurate and reliable, and can truly reflect the morphological characteristics of remote-sensing image.

In order to verify the segmentation precision of our proposed algorithm in the real complicated database, we adopt the high resolution remote-sensing image of the original Washington area. According to the comparison results in figure 3, we can see that compared with other comparison algorithms in the segmentation of the real complicated sample image, our proposed algorithm has relatively good segmentation precision. Through several segmentation results in the same dataset, we can see that the segmentation precision does not have large fluctuation compared with other algorithm, and the segmentation precision of our proposed algorithm is between 92.8%~94.2%, the segmentation precision of the algorithm proposed by JSEG algorithm is between 80.9%~86.7%. Theoretical analysis and computer simulation results show that our proposed algorithm can effectively use a large number of unlabeled samples and can improve the segmentation accuracy.

Segmentation experiment results are shown as Fig. 6.

According to the above segmentation experiment results, it is obvious that regional boundary has more accurate segmentation result after $H$ value is adopted to correct the local $J$ value. Through some detailed comparison, it is found that our proposed algorithm effectively solves the problem of inaccurate segmentation.
boundary, but the algorithm also has over-segmentation phenomenon in the same region. For example, the bleacher on the playground in Washington image exists obvious over-segmentation phenomenon whether in the traditional algorithm or in the improved algorithm.

V. CONCLUSION

Aiming at the flaws of JSEG algorithm and the color and texture feature of remote-sensing image in the process of image segmentation, a color remote-sensing image segmentation based on improved region filter is proposed in this paper. In order to well describe the color homogeneity in region, the local homogeneity matrix can be adopted to correct the local value in traditional JSEG segmentation algorithm so as to accurately reflect the boundary region and enhance the accuracy of image segmentation. Simulation experiment results demonstrate that our proposed algorithm can efficiently overcome the inaccurate segmentation of boundary region. In addition, compared with the traditional filter algorithm, our proposed algorithm can extract weak edge more accurately so as to achieve the better segmentation result in high resolution remote-sensing image segmentation.

REFERENCES


© 2014 ACADEMY PUBLISHER
Instructions for Authors

Manuscript Submission

All paper submissions will be handled electronically in EDAS via the JMM Submission Page (URL: http://edas.info/newPaper.php?c=7325). After login EDAS, you will first register the paper. Afterwards, you will be able to add authors and submit the manuscript (file). If you do not have an EDAS account, you can obtain one. Along with the submission, Authors should select up to 3 topics from the EDICS (URL: http://www.academypublisher.com/jmm/edmics.html), and clearly state them during the registration of the submission.

JMM invites original, previously unpublished, research papers, review, survey and tutorial papers, application papers, plus case studies, short research notes and letters, on both applied and theoretical aspects. Submission implies that the manuscript has not been published previously, and is not currently submitted for publication elsewhere. Submission also implies that the corresponding author has consent of all authors. Upon acceptance for publication transfer of copyright will be made to Academy Publisher, article submission implies author agreement with this policy. Manuscripts should be written in English. Paper submissions are accepted only in PDF. Other formats will not be accepted. Papers should be formatted into A4-size (8.27" x 11.69") pages, with main text of 10-point Times New Roman, in single-spaced two-column format. Authors are advised to follow the format of the final version at this stage. All the papers, except survey, should ideally not exceed 12,000 words (14 pages) in length. Whenever applicable, submissions must include the following elements: title, authors, affiliations, contacts, abstract, index terms, introduction, main text, conclusions, appendixes, acknowledgement, references, and biographies.

Conference Version

Submissions previously published in conference proceedings are eligible for consideration provided that the author informs the Editors at the time of submission and that the submission has undergone substantial revision. In the new submission, authors are required to cite the previous publication and very clearly indicate how the new submission offers substantively novel or different contributions beyond those of the previously published work. The appropriate way to indicate that your paper has been revised substantially is for the new paper to have a new title. Author should supply a copy of the previous version to the Editor, and provide a brief description of the differences between the submitted manuscript and the previous version.

If the authors provide a previously published conference submission, Editors will check the submission to determine whether there has been sufficient new material added to warrant publication in the Journal. The Academy Publisher’s guidelines are that the submission should contain a significant amount of new material, that is, material that has not been published elsewhere. New results are not required; however, the submission should contain expansions of key ideas, examples, elaborations, and so on, of the conference submission. The paper submitting to the journal should differ from the previously published material by at least 30 percent.

Review Process

Submissions are accepted for review with the understanding that the same work has been neither submitted to, nor published in, another publication. Concurrent submission to other publications will result in immediate rejection of the submission.

All manuscripts will be subject to a well established, fair, unbiased peer review and refereeing procedure, and are considered on the basis of their significance, novelty and usefulness to the Journals readership. The reviewing structure will always ensure the anonymity of the referees. The review output will be one of the following decisions: Accept, Accept with minor changes, Accept with major changes, or Reject.

The review process may take approximately three months to be completed. Should authors be requested by the editor to revise the text, the revised version should be submitted within three months for a major revision or one month for a minor revision. Authors who need more time are kindly requested to contact the Editor. The Editor reserves the right to reject a paper if it does not meet the aims and scope of the journal, it is not technically sound, it is not revised satisfactorily, or if it is inadequate in presentation.

Revised and Final Version Submission

Revised version should follow the same requirements as for the final version to format the paper, plus a short summary about the modifications authors have made and author's response to reviewer's comments.

Authors are requested to use the Academy Publisher Journal Style for preparing the final camera-ready version. A template in PDF and an MS word template can be downloaded from the website. Authors are requested to strictly follow the guidelines specified in the templates. Only PDF format is acceptable. The PDF document should be sent as an open file, i.e. without any data protection. Authors should submit their paper electronically through email to the Journal's submission address. Please always refer to the paper ID in the submissions and any further enquiries.

Copyright

Submission of your paper to this journal implies that the paper is not under submission for publication elsewhere. Material which has been previously copyrighted, published, or accepted for publication will not be considered for publication in this journal. Submission of a manuscript is interpreted as a statement of certification that no part of the manuscript is copyrighted by any other publisher nor is under review by any other formal publication.

Submitted papers are assumed to contain no proprietary material unprotected by patent or patent application; responsibility for technical content and for protection of proprietary material rests solely with the author(s) and their organizations and is not the responsibility of the Academy Publisher or its editorial staff. The main author is responsible for ensuring that the article has been seen and approved by all the other authors. It is the responsibility of the author to obtain all necessary copyright release permissions for the use of any copyrighted materials in the manuscript prior to the submission. More information about permission request can be found at the web site.

Publication Charges and Re-print

The author's company or institution will be requested to pay a flat publication fee of EUR 360 for an accepted manuscript regardless of the length of the paper. The page charges are mandatory. Authors are entitled to a 30% discount on the journal, which is EUR 100 per copy. Reprints of the paper can be ordered with a price of EUR 100 per 20 copies. An allowance of 50% discount may be granted for individuals without a host institution and from less developed countries, upon application. Such application however will be handled case by case.
